

Tesis Doctoral

# Modelos de baja dimensión para canto de aves y aplicación a interfaces cerebro-máquina

Arneodo, Ezequiel Matías

2012

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Arneodo, Ezequiel Matías. (2012). Modelos de baja dimensión para canto de aves y aplicación a interfaces cerebro-máquina. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Arneodo, Ezequiel Matías. "Modelos de baja dimensión para canto de aves y aplicación a interfaces cerebro-máquina". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2012.

**EXACTAS** UBA

Facultad de Ciencias Exactas y Naturales



**UBA**

Universidad de Buenos Aires



Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales  
Departamento de Física

## Modelos de baja dimensión para canto de aves y aplicación a interfaces cerebro-máquina.

Tesis presentada para optar al título de Doctor de la  
Universidad de Buenos Aires en el área Ciencias Físicas

Ezequiel Matías Arneodo

**Director de Tesis:** Gabriel B. Mindlin.

**Consejero de estudios:** Mariano Sigman.

**Lugar de trabajo:** Laboratorio de Sistemas Dinámicos, Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

Buenos Aires, febrero de 2012



---

## Resumen

El canto de aves oscinas es un comportamiento complejo aprendido que surge de la interacción entre un sistema nervioso central y una serie de dispositivos periféricos bio-mecánicos. Es una tarea intensa la de dilucidar cuánta de la complejidad es debida a la enorme dimensión de la dinámica del sistema nervioso, y cuánta a la de la respuesta mecánica de la periferia. Este último punto resulta de particular interés para la física, porque los sistemas bio-miméticos empleados por los seres vivos son abrumadoramente no lineales, y capaces de presentar comportamiento complejo aún ante instrucciones sencillas. Esta tesis se propone mostrar que modelos vocales sumamente sencillos son capaces de reproducir las sutilezas presentes en el canto de las aves oscinas, siendo la sencillez del modelo de tal magnitud que su solución es computable mediante un procesador de señales digitales (DSP). Esto permite diseñar una nueva estrategia para la bio-mimética prostética vocal: la integración en tiempo real de modelos sencillos, controlados por pocas señales fisiológicas. En esta tesis se muestra la plausibilidad de esta estrategia construyendo un prototipo de ese dispositivo.

**Palabras clave:** canto de aves - control motor - dinámica no-lineal - diamante mandarín - bio-prostética - ecuaciones diferenciales con retardo

---

## Low-Dimensional models for Birdsong and application to Brain Machine Interfaces

### Abstract

The complex vocalizations composing birdsong emerge from the interaction between the nervous system and a nonlinear peripheral device, its vocal organ. It is an area of intensive work to elucidate how much of the complexity of it is due to the huge dimensionality of the dynamics of the nervous system, and how much due to the mechanical response of the periphery. This last point is of particular interest for applied physics, as bio-mechanic systems used by living things are overwhelmingly non-linear, and capable of presenting complex behaviors even as a response to simple instructions. We present a bio-prosthetic device based on a model for the zebra finch vocal organ, which is capable of producing the subtleties present in the actual song of the birds. This model is simple to the extent that its solution can be computed by digital signal processors (DSP), when fed with actual physiological motor instructions by a freely behaving subject. This allows for the design of a new strategy for biomimetic prosthesis: real-time integration of simple models, controlled by few physiological signals. In this thesis we show the plausibility of this strategy by building a prototype of such a device.

**Keywords:** birdsong - motor control - nonlinear dynamics - zebra finch - bio-prosthetic - delay differential equations

## Índice general

Resumen	I
Abstract	II
<b>Parte 1. Intro</b>	<b>1</b>
Capítulo 1. Introducción	3
1. El canto de las aves oscinas	3
2. Fuentes de complejidad	3
3. La importancia de contar con un modelo	5
4. Esta tesis	6
Capítulo 2. Un modelo animal: el canto de los oscinos	11
1. El modelo animal	11
2. Modelo de aleteo extendido	12
<b>Parte 2. Interacciones en la periferia: acople fuente filtro</b>	<b>15</b>
Capítulo 3. Buscando complejidad en interacciones sencillas	17
1. Fuentes de complejidad	17
2. Observaciones experimentales sugestivas	18
3. Un modelo no tan simple de acople fuente-tracto	18
4. La inspiración en los oscinos: modelo de fuente sencilla acoplada al tracto	20
5. Un tracto simple y acoplado	22
6. La complejidad de la dimensión infinita: saltos de frecuencia	23
Capítulo 4. La dinámica de la frecuencia: disipación infinita	25
1. La aproximación	25
2. Dinámica de los ciclos límite	26
3. Origen dinámico de los saltos de frecuencia	30
Capítulo 5. La dinámica de la amplitud: la fuente <i>cuasi</i> -lineal	33
1. Motivación	33
2. La aproximación	35
3. Efectos dinámicos observables en la amplitud	38
4. Propiedades acústicas de las soluciones	39
5. Discusión	43
<b>Parte 3. Interfaces cerebro-máquina basados en los modelos</b>	<b>45</b>
Capítulo 6. Comportamiento motor complejo: el canto del diamante Mandarín	47

---

1. Canto del diamante Mandarín	47
2. Modelo del órgano fonador	49
3. Una restricción de la biomecánica	51
Capítulo 7. Síntesis de canto en tiempo real	55
1. Integración en tiempo real mediante DSP	55
2. Simplificando lo simple	56
3. Canto sintético en tiempo real	59
Capítulo 8. Interfaz cerebro-máquina para diamante mandarín	65
1. BMI con modelo de la biomecánica	65
2. Un tracto vocal superior más completo	66
3. Implementación	68
4. Perspectivas para BMI	76
<b>Parte 4. Desenlace y final</b>	<b>77</b>
Capítulo 9. Conclusiones	79
1. Objetivos cumplidos	79
2. Aprendimos algo, hoy	82
Capítulo 10. Perspectivas	85
1. Todavía más complejidad en la periferia	85
2. Siringe electrónica	85
3. BMI	87
4. Jugar en toda la cancha	89
Apéndice A. Circuito para rectificación y suavizado de EMG	91
1. Rectificación y amplificación	91
2. Implementación electrónica analógica	91
Apéndice B. Amplificador de la señal de presión	93
1. Diseño del Amplificador	93
Apéndice. Encerar, pulir	97
Apéndice. Bibliografía	99

**Parte 1**

**Intro**





## Introducción

### 1. El canto de las aves oscinas

**1.1. Aprendizaje de comportamiento vocal complejo.** El canto de las aves oscinas es un comportamiento motor complejo que se aprende, durante un período sensible. Esto, que es compartido con los humanos y con otras pocas especies animales, y los paralelos encontrados entre los mecanismos que originan sonido en su órgano fonador y en el de los humanos, las hacen un modelo animal ideal para estudiar las interacciones entre un sistema nervioso central y una periferia biomecánica que conducen al aprendizaje de un comportamiento motor no trivial [ZM04].

Las aves que pertenecen al suborden de los oscinos (llamadas también “aves canoras”) al igual que los humanos, adquieren sus vocalizaciones luego de un proceso de aprendizaje. Este requiere que se escuchen las vocalizaciones de un tutor durante un periodo de memorización y que se escuchen luego durante un período de práctica. A lo largo de la primera etapa (sensorial), se forma una representación interna de las vocalizaciones de los tutores. Durante la segunda etapa (sensorimotora), esta es comparada con su propia producción, que es corregida hasta lograr una copia estereotipada del canto de la especie [ZM04].

Las vocalizaciones que pueden producir las aves canoras son complejas. Si bien la complejidad de los sonidos y la estructura temporal varía entre especies, un amplio rango de características acústicas es cubierto, rango que abarca una diversidad de contenidos espectrales, frecuencias fundamentales, y efectos acústicos no lineales como bi-fonación (dos frecuencias no relacionadas), transiciones rápidas entre frecuencias, y sonidos ruidosos semejantes al caos determinista [ZS04, ZRS08].

### 2. Fuentes de complejidad

**2.1. El aparato fonador.** El comportamiento emerge de la interacción entre dos sistemas. El sistema nervioso central genera instrucciones motoras con las que se conduce una periferia biomecánica, el órgano fonador. Para generar las instrucciones motoras es necesaria la interacción entre una serie de núcleos neuronales, que conforman el sistema de canto [ZM04]. La salida de este sistema controla una fuente de sonido bipartita (la siringe) y el sistema respiratorio.

El órgano fonador del ave incluye a la siringe, al sistema respiratorio, y al tracto vocal superior [ML05]. En la siringe de los oscinos, dos conjuntos de membranas ubicadas en la juntura entre los bronquios y la tráquea modulan el flujo de aire y producen sonido. Cada una de las fuentes puede ser controlada mediante conjuntos poco numerosos de músculos independientes

---

[GC04]. Los sonidos producidos en la siringe atraviesan luego el tracto vocal superior [Now87, ML05, AM09, APM11]. En el tracto vocal superior, el sonido producido por la fuente es alterado. Si no existe interacción acústica entre la fuente y el tracto, este último actúa como un filtro, modificando el peso relativo de las componentes espectrales del canto.

Las no linealidades del aparato fonador pueden introducir complejidad y características acústicas variadas en los sonidos, incluso al ser él conducido por patrones motores sencillos [GCM01, LGM02]. Mediante gestos suficientemente precisos como para explotar el control independiente de las dos fuentes de sonido en la siringe, pueden producirse vocalizaciones que presenten una estructura espectral o temporal compleja [ZS04].

Las interacciones acústicas o mecánicas entre los dos lados de la siringe o con el tracto vocal han sido estudiadas como posibles mecanismos subyacentes a fenómenos no lineales observados en el canto de algunas especies, como el ruiseñor (*Mimus polyglottos*) [ZRS08]. Sugestivamente, se ha observado además que la actividad del tracto se coordina con la de la siringe de manera que sus resonancias están en sintonía con la frecuencia fundamental de las vocalizaciones [RSFB06, BPDH97].

**2.2. Dinámica no lineal en la periferia.** La neurociencia del canto de las aves ha dedicado mucho esfuerzo a comprender cómo se producen los gestos motores en el sistema nervioso central, y cuáles son los mecanismos neuronales de adquisición y mantención [ZM04, GC04, SM02, YM96]. En esta tesis nos dedicamos al estudio de la periferia biomecánica. Nos proponemos contribuir a la pregunta de cuánta de la complejidad del comportamiento tiene su origen en la dinámica altamente no lineal de la fuente. Esperamos ser capaces además de allanar el camino a experimentos en los que se pueda ir aún más lejos, y determinar de qué modo el paquete de soluciones posibles de la periferia restringe las posibilidades del sistema nervioso central a la hora de adquirir y mantener los gestos motores con los que la conduce.

La periferia es un sistema biomecánico que responde a instrucciones fisiológicas engendradas en el sistema nervioso central. La biomecánica puede ser modelada matemáticamente con sistemas de ecuaciones diferenciales, y en las soluciones de esos sistemas pueden ser identificados observables del resultado: el comportamiento emergente [LGM02, ML05]. Se busca que los parámetros de los modelos sean fisiológicamente relevantes, de modo que las instrucciones motoras determinen las soluciones y el comportamiento de los sistemas al determinar las regiones del espacio de parámetros en las que un sistema se encuentra.

Utilizando procedimientos estándar de la dinámica no lineal, es posible explorar los modelos analíticamente. De este modo buscamos determinar los mecanismos que subyacen a la generación del comportamiento complejo, de modo robusto y general. Es decir, es posible hacer modelos que expliquen el origen del comportamiento que sean predictivos y cuantitativos.

Haciendo modelos, explorando la relación de los parámetros del modelo con las instrucciones motoras y de las soluciones con los observables (el comportamiento), podemos contribuir a la pregunta de dónde se origina la complejidad. Podemos, con base en esos modelos y en su implementación

---

en dispositivos que emulen el comportamiento a partir de las instrucciones, diseñar experimentos que permitan estudiar el rol de la periferia en la generación de la complejidad y de qué modo su presencia, su ausencia y la alteración de su funcionamiento condicionan al sistema generador de patrones.

### 3. La importancia de contar con un modelo

**3.1. Experimentos de realimentación sensorial alterada.** La producción del habla en humanos y del canto en oscinos depende de la retroalimentación sensorial para el aprendizaje y la mantención de gestos motores estereotipados. En aves, un circuito neuronal específico es esencial en la modificación del canto durante la ontogenia o en adultos, cuya realimentación auditiva se ha interrumpido [BMA84, BNN96]. Tras el enmudecimiento o enmascaramiento de la realimentación usando ruido blanco, el canto de los adultos tiende a deteriorarse a un ritmo que depende de la especie en la que se hace el experimento [OY97, LK99, SB06]. El canto más estudiado (el del diamante mandarín, *Taeniopygia guttata*) consiste de una serie altamente estereotipada de 4 a 8 sílabas, compuestas de sonidos tonales, sonidos espectralmente ricos, y elementos más ruidosos [ZM04]. Las deficiencias en las vocalizaciones, para esta especie, ocurren en adultos luego de días a semanas de interrupción crónica de la realimentación. Los mecanismos centrales de producción de canto han sido exitosamente estudiados, en términos de qué núcleos neuronales están involucrados [HFKF02, FKH04, YM96], pero no se ha atacado hasta ahora directamente la pregunta de cuál es la naturaleza de las instrucciones motoras específicas de los distintos elementos de sonido. Por ende, hasta aquí los experimentos de realimentación sensorial alterada han sido realizados mediante alteraciones de la realimentación que no guardan una relación sistemática con las instrucciones motoras que el ave genera para producirlas.

En esta tesis presentamos la construcción de una siringe electrónica, un dispositivo que, a partir de señales fisiológicas, integra los modelos del aparato fonador en tiempo real [ZMP<sup>+</sup>05, SAGM08]. Éste hará posible la realización de experimentos de realimentación alterada, por medio de los cuales se podrá disociar el rol que juega la retroalimentación en el aprendizaje y la mantención del canto en el diamante mandarín. En estos experimentos, las propiedades acústicas de las vocalizaciones sintéticas son controladas por el ave mediante sus propias instrucciones motoras. Así, las alteraciones en la realimentación pueden ser hechas de modo consistente con alteraciones en las instrucciones motoras.

**3.2. BMI basada en modelos de baja dimensión.** La estrategia habitual de BCI y BMI (*Brain Computer Interfaces y Brain Machine Interfaces*) consiste en descifrar los comandos motores a partir de registros de actividad fisiológica en el cerebro y utilizar esta actividad para controlar dispositivos bio-miméticos [SCWM06, CWR<sup>+</sup>08, CLC<sup>+</sup>03]. En [CLC<sup>+</sup>03], por ejemplo, una cantidad de registros de la actividad de decenas a cientos de neuronas en diferentes áreas de la corteza de primates se utilizan para accionar un brazo robótico. En un trabajo reciente, Cichocki *et. al.* discuten las perspectivas de la utilización de electroencefalogramas (EEG) para generar

---

soluciones BCI no invasivas [CWR<sup>+</sup>08]. En estos ejemplos (como en muchas otras implementaciones de BCI), el problema crucial es la clasificación de las características de la gran cantidad de datos que corresponden a un conjunto determinado de tareas motoras. La extracción de características se realiza mediante diferentes técnicas que incluyen la descomposición lineal en una diversidad de espacios vectoriales, combinadas con algoritmos de aprendizaje automático [SCWM06, PS00, CWR<sup>+</sup>08]. De esta manera, se logra un control preciso de efectores bio-miméticos para un número finito de tareas específicas, como agarrar o mover un cursor. Debido a que la extracción de características y la selección se realiza sobre bases estadísticas en un espacio de parámetros de alta dimensión, se requiere cálculo pesado. Esto impone una restricción a la portabilidad de estas soluciones y a la complejidad de la tarea a realizar.

Nuestra comprensión actual de la biofísica del órgano vocal aviar en términos de un modelo matemático de baja dimensión, nos permite proponer un ejemplo de un tipo diferente de solución bio-protésica. El aporte viene, sobre todo, de nuestra capacidad para identificar los mecanismos por los cuales el comportamiento dinámico complejo se produce cuando los sistemas periféricos son impulsados por instrucciones suaves contenidas en un número reducido de registros fisiológicos. El modelo predice una diversidad de soluciones cualitativamente diferentes al sistema para trayectorias continuas en un espacio de parámetros. Este espacio de parámetros sugerido por el modelo es fisiológicamente pertinente.

#### 4. Esta tesis

Esta tesis se dedica a entender de qué modos puede un efector biomecánico periférico ser responsable de originar comportamiento motor complejo. Partimos de lo avanzado en la descripción de los mecanismos de control de la siringe y de observaciones sistemáticas de la actividad de la periferia correlacionadas con observables del canto. Con esta base, extendemos la idea de que parte de la complejidad surge de la no-linealidad de la fuente de sonido, proponiendo un modelo matemático de interacción entre la siringe y el tracto vocal superior.

Una propiedad de los sistemas no lineales es que soluciones cualitativamente distintas surgen de caminos suaves en un espacio de parámetros. En esta tesis evaluamos la propuesta de que la comprensión de la dinámica no lineal de la periferia permite construir soluciones bio-miméticas de bajo costo computacional.

La estructura de la presentación, y una síntesis de lo que se muestra en cada parte, se encuentran a continuación.

**Parte 1: Capítulos introductorios.** Esta parte finaliza con un capítulo que se dedica a introducción del modelo matemático que describe la física de la siringe bipartita de los oscinos. Presentamos el modelo de aleteo que presentó Titze para dar cuenta del inicio de oscilaciones en las cuerdas vocales humanas, y el modelo extendido, que describe adecuadamente las oscilaciones en los tejidos oscilantes en la siringe.

---

**Capítulo 2: Un modelo animal: el canto de los oscinos.** se introduce el caso de las aves oscinas, poniendo énfasis en lo que se conoce en la actualidad acerca de cuáles son los mecanismos de producción de sonido en la periferia biomecánica y de control de las propiedades acústicas de las vocalizaciones.

Porque existen paralelos con los humanos en la forma en la que los oscinos aprenden y en la biofísica de su aparato fonador, tomamos este modelo como un caso de estudio al investigar las capacidades que tiene un sistema periférico biomecánico complejo para producir diversidad en el comportamiento aún conducido por instrucciones fisiológicas motoras sencillas. Explotando esta capacidad, este modelo animal será también el utilizado para proponer un ejemplo de dispositivo bio-prostético, en el que el modelado matemático de baja dimensión, en vez de las técnicas estadísticas, será la guía para identificar registros fisiológicos y patrones adecuados para conducir la interfase sintética entre el patrón motor y el comportamiento.

En la segunda y tercera parte de la tesis se describe el trabajo realizado.

**Parte 2: Interacciones en la periferia: acople fuente filtro.** En la segunda parte, extendemos la búsqueda de complejidad en la periferia a las interacciones acústicas entre la fuente de sonido y el tracto vocal superior.

**Capítulo 3: Buscando complejidad en interacciones sencillas.** En el plan de determinar qué tan capaz es la periferia biomecánica de introducir complejidad en el comportamiento, resultan ineludibles las interacciones acústicas entre la fuente de sonido y el tracto vocal.

Proponemos un modelo matemático simple para la fuente, para poder identificar los fenómenos que se agregan al introducir el acople con el tracto. Si la válvula labial está acoplada a un tubo (con el que aproximamos al tracto), su dinámica se verá afectada por las reflexiones de la onda de sonido a la salida, con un el retardo que le implica a la onda de sonido viajar hasta la salida y reflejarse de vuelta.

El sistema resultante al introducir la interacción es un sistema de ecuaciones diferenciales con retardo (*DDE*). El término con retardo dependerá del estado del sistema en un tiempo anterior. De este modo, conocer el futuro del estado del sistema para un instante dado, requiere como condiciones iniciales a los estados del sistema en un intervalo de tiempo previo. La dinámica resultante es la de un sistema de dimensión infinita. Para estudiar este sistema y determinar los orígenes dinámicos de los cambios en el modo de oscilación que dan origen a una variedad de propiedades acústicas de las soluciones, reduciremos el sistema utilizando ciertas aproximaciones.

**Capítulo 4: La dinámica de la frecuencia: disipación infinita.** Uno de los efectos candidatos a tener su origen en complejidad introducida por el acople acústico entre la fuente y el tracto es la ocurrencia de saltos de frecuencia en las vocalizaciones, no relacionados con variaciones repentinas en las instrucciones motoras registradas (presión en los sacos aéreos, flujo de aire a través de la siringe) [ZRS08]. En sistemas

---

complejos, este tipo de saltos puede tener su origen en la histéresis de ciclos límite: un salto de frecuencia puede producirse al atravesar, en el espacio de parámetros adecuado, una región de coexistencia de soluciones oscilatorias estables.

La dinámica de las soluciones periódicas de un sistema de ecuaciones diferenciales con retardo puede ser explorada analíticamente, si se lleva al sistema al límite de alta disipación [EG08]. En ese límite, podremos encontrar ecuaciones de fase para el sistema cuando no hay acople. Estas ecuaciones describen la dinámica de las trayectorias en el ciclo límite utilizando una única dimensión, porque la disipación es lo suficientemente grande como para que las perturbaciones a la amplitud del ciclo sean pequeñas y decaigan rápidamente. Luego, al introducir el acople con el filtro como una perturbación a la dinámica de la fase, nos será posible estudiar analíticamente las bifurcaciones que den origen a coexistencia de soluciones periódicas estables.

**Capítulo 5: La dinámica de la amplitud: la fuente cuasi-lineal.** El observable más común en el estudio del canto de los oscinos es la presión acústica (la amplitud del sonido). Si el acople fuente-tracto introduce efectos más sutiles que los saltos de frecuencia registrados en algunas especies, esperamos que estos sean identificables en ese observable. Observamos numéricamente en nuestro modelo que las resonancias de un tracto que actúa como un filtro pasivo son cualitativamente distintas a las de un sistema fuente-tracto interactuantes.

Recurrimos a la aproximación débilmente no lineal para encontrar los orígenes dinámicos del fenómeno. En esta aproximación, la dinámica del sistema fuente-tracto es vista como una perturbación a la dinámica de un oscilador armónico. En estas condiciones, obtenemos expresiones analíticas que describen las desviaciones de las soluciones del sistema de oscilaciones armónicas.

Consecuentemente, las bifurcaciones que conducen a soluciones oscilatorias no triviales son identificables. Al proponer el estudio del sistema en una aproximación débilmente no lineal, nos involucramos con las propiedades no lineales del sistema fuente-tracto interactuantes que hacen mella en propiedades acústicas evidentes en la amplitud de las vocalizaciones.

### **Parte 3: Interfaces cerebro-máquina basados en los modelos.**

El desarrollo de modelos de baja dimensión que dan cuenta de la dinámica no lineal de la periferia biomecánica plantea un modo de construir interfaces sintéticas entre cerebro y comportamiento. Por un lado, porque el nivel de descripción adecuado permite identificar un número reducido de parámetros fisiológicamente relevantes, sin necesidad de recurrir a procedimientos estadísticos en un gran conjunto de registros neurofisiológicos como EEG (*Electroencephalography*) o fMRI (*Functional magnetic resonance imaging*). Por otro lado, porque en estos modelos el comportamiento se obtiene al integrar un sistema de unas pocas ecuaciones diferenciales ordinarias a partir de caminos sencillos en el espacio de los parámetros.

---

En la tercera parte de la tesis se presenta el desarrollo de un prototipo de aparato fonador prostético para el diamante mandarín, que es manejado por un sujeto comportándose libremente mediante un gesto motor fisiológico.

**Capítulo 6: Comportamiento motor complejo: el canto del Diamante Mandarín.** Uno de los cantos más estudiados es el del diamante mandarín. Presentamos un modelo que fue desarrollado en [SAGM08, AM08], en el que pueden ocurrir oscilaciones a través de dos bifurcaciones distintas. Las propiedades de esas oscilaciones se corresponden con las propiedades acústicas de los diversos sonidos que produce el ave.

**Capítulo 7: Síntesis de canto en tiempo real.** En el avance hacia la construcción de un dispositivo bio-prostético que emule el comportamiento del órgano fonador aviar, exploramos la factibilidad de la implementación. Para ello simplificamos computacionalmente el modelo manteniendo las soluciones cualitativamente distintas y los mecanismos dinámicos que originan oscilaciones, reduciéndolo a una forma normal. Encontramos un conjunto de transformaciones de los gestos motores (presión en el saco aéreo y actividad del músculo siríngeo ventral) a parámetros de la forma normal con el que, mediante integración numérica en un procesador de señales digitales, se reproduce canto sintético comparable con el real.

**Capítulo 8: Interfaz cerebro-máquina para diamante mandarín.** En última instancia, convergemos a que el contenido de esta tesis aporta confianza a la ambición de generar una clase de dispositivos controlados por el cerebro. En los paradigmas actuales, se realiza estadística sobre las señales fisiológicas, y se encuentran las relaciones entre patrones y el comportamiento esperado (la realimentación sensorial buscada por el sujeto). [SCWM06, Nic01, LCiA04]. Nuestra propuesta se basa en modelos predictivos de la respuesta de la periferia ante instrucciones motoras contenidas en registros fisiológicos. En este capítulo desarrollamos un prototipo de este dispositivo, en el que un ave previamente enmudecida conduce la integración del modelo en tiempo real, mediante el gesto motor registrado a través de la presión en su saco aéreo.

**4.1. Parte 4: Desenlace y final.** En esta parte están el cierre de la tesis, las conclusiones, lo que se aprendió y los caminos que quedan planteados, a partir de los éxitos parciales de lo que se desarrolló hasta acá.





## Un modelo animal: el canto de los oscinos

El canto de las aves oscinas presenta un conjunto diverso de características acústicas. Mucha de esa diversidad puede ser explicada por la dinámica que exhibe un modelo matemático de baja dimensión para la fuente de sonido que forman las membranas que oscilan en la siringe [LGM02, SAGM08].

### 1. El modelo animal

**1.1. La siringe.** La siringe de las aves oscinas está ubicada en el saco aéreo interclavicular, en la juntura entre los bronquios y la tráquea. Está compuesta por estructuras cartilaginosas, masas de tejido conectivo, membranas y músculos.

La siringe tráqueo-bronquial de los oscinos es una estructura simétrica bilateral con dos fuentes de sonido. En cada lado de la siringe, hay masas de tejido vibrante que forman una válvula (figura 1.1a). La parte lateral de cada válvula está compuesta por el *labium lateral* (LL). A la parte medial la componen el *labium medial* (ML) en su parte anterior y la *membrana timpaniforme medial* (MTM) en la posterior.

**1.2. Mecanismo de producción de sonido.** El mecanismo de fonación de las aves es similar al de los humanos [ZM04, ML05]. La presión del saco aéreo establece un flujo de aire a través de la siringe, que funciona como una válvula de labios vibrantes. Bajo ciertas condiciones en la presión ejercida por el flujo de aire y la geometría de la válvula, las membranas realizan un movimiento oscilatorio sostenido. Estas oscilaciones modulan el flujo de aire y producen una onda de presión acústica. Esta onda de sonido ingresa en el tracto vocal superior. Si no existe interacción acústica entre la válvula siríngea y el tracto, el efecto del tracto vocal es el de un filtro que resalta las frecuencias cercanas a la resonancia y atenúa el resto.

En la siringe traqueo-bronquial de los oscinos, las válvulas se forman al contraerse el lumen siríngeo y aproximarse entre sí los labios mediales y laterales (ML y LL). Este gesto muscular lleva a la siringe a la posición fonatoria y es seguido de un aumento de la presión sub-siríngea. El flujo de aire generado a través de la válvula formada por los labios induce bajo ciertas condiciones oscilaciones labiales [aSGP99].

**1.3. La actividad muscular.** En la figura 1.1b se encuentra un esquema de la disposición de los músculos de la siringe. La actividad de los músculos *siringealis dorsalis* (dS) y *tracheo-bronquialis dorsalis* (dTb) produce un cierre activo del lumen siríngeo. La apertura activa del lumen está a cargo de los músculos *tracheo-lateralis* (TL) y *tracheo-bronchialis ventralis* (vTB) [GS96].

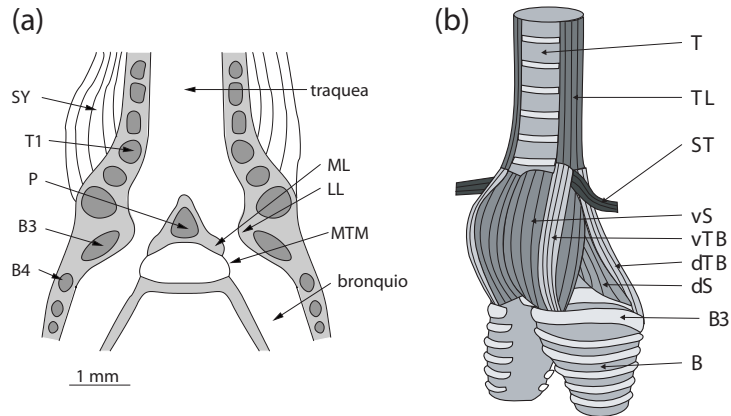


FIGURA 1. Esquema de la siringe de un ave oscina. (a) Sección ventral. Se forman dos válvulas labiales por aducción de los labios laterales y mediales de cada lado. LL, ML: labios laterales y mediales; MTM: membranas timpaniformes mediales; SY: músculo siringeo. (b) Vista ventro-lateral. La frecuencia fundamental de las vocalizaciones está correlacionada con la actividad de los músculos *siringealis ventralis* (vS). TL: m. *tracheo-lateralis*; ST: m. *sternotrachealis*; vTB: m. *tracheo-bronchialis ventralis*; dTB: m. *tracheo-bronchialis dorsalis*; dS: m. *siringealis dorsalis*; T: tráquea; B: bronquio. (extraído de [ML05]).

Durante la vocalización, la actividad del músculo *siringealis ventralis* (vS) está fuertemente correlacionada con la frecuencia fundamental de la vocalización [GS96, GC04]. El músculo *sternotrachealis* (ST) no ha sido identificado con un rol diferenciado.

## 2. Modelo de aleteo extendido

Un modelo para explicar el mecanismo de producción de sonido en la siringe de las aves que presenta un buen compromiso entre el nivel de descripción y la complejidad computacional se presentó en [GCM01, LGM02]. Se basa en el mecanismo de “aleteo” de Titze, uno de los modelos más sencillos para dar cuenta de la transferencia de la energía cinética del aire a las oscilaciones de las cuerdas vocales en seres humanos [Tit88]. Sobre la base de observaciones experimentales, se asume que las piezas de tejido blando o labios que forman la válvula vibrante son soporte de dos modos de vibración: una onda superficial propagándose en dirección vertical, y una oscilación lateral de sus posiciones centrales. Es decir, los labios realizan un movimiento de “aleteo”. Ha sido sugerido por la videografía de las cuerdas durante la fonación [Tit88, LG99, FSPM98] que estos modos están coordinados de manera que se extrae energía de del flujo de aire en cada ciclo, haciendo posibles oscilaciones sostenidas.

La descripción cinemática de este mecanismo se realiza en términos del desplazamiento del punto medio de cada labio  $x$  respecto de su posición de equilibrio. Suponemos que el perfil de la válvula siringea es trapezoidal,

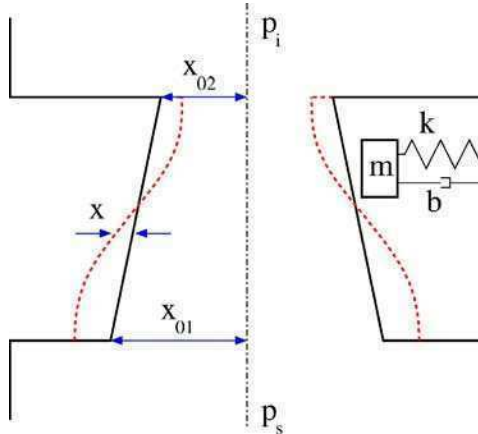


FIGURA 2. Esquema de la sección frontal de la válvula siríngica. La variable  $x$  es el desplazamiento del punto medio del labio respecto de la posición de equilibrio (prefonatoria).

como se muestra en la figura 2. El desplazamiento de la posición del punto medio de cada labio obedece la segunda ley de Newton de un oscilador no lineal forzado de masa  $m$ ,

$$\begin{cases} \frac{dx}{dt} = y, \\ \frac{dy}{dt} = (1/m) [-k(x)x - b(y)y - cx^2y + a_l p_{av}], \end{cases} \quad (1)$$

donde  $a_l$  es el área media del labio. En el lado derecho de la segunda ecuación, el primer término describe la restitución no lineal elástica de los labios, el segundo término representa la disipación no lineal, y el tercero una saturación no lineal que limita el movimiento labial. El sistema es forzado por el último término, que representa la fuerza ejercida por la presión inter-labial.

En el término de forzado,  $p_{av}$  es el promedio espacial de la presión de conducción. Es una función de la geometría de las membranas que depende de la presión sub-siríngica  $p_s$  y de la relación de las secciones traqueal y bronquial de la válvula labial. Dado que la contribución de la presión a la entrada del tracto  $p_i$  a la presión inter-labial es mucho menor que  $p_s$ , se desprecia [AM09]. Esto significa que no se considera que la fuente de sonido y el tracto vocal interactúen acústicamente.

La coordinación de los dos modos de oscilación soportados por los tejidos, observada experimentalmente como una diferencia de fase entre los extremos superior e inferior de los labios durante la fonación [LG99, FSPM98], se tiene en cuenta en esta función. Mediante la introducción de un parámetro fenomenológico  $\tau_l$  que describe el tiempo que tarda la onda que se propaga hacia arriba en los labios para cubrir la mitad de su longitud, las áreas bronquial y traqueal de la válvula ( $a_1$  y  $a_2$ ) llegan a ser escritas en términos de esta constante de tiempo, del desplazamiento del punto medio de los labios y de las posiciones de reposo (pre-fonación) de los extremos de los labios ( $x_{01}$  y  $x_{02}$ ) [Tit88].

Finalmente, una versión fenomenológicamente corregida de la ecuación de Bernoulli se utiliza para escribir la presión media a través de la válvula

---

en términos de la relación de las secciones y la presión sub-siríngea [Tit88, LGM02]:

$$p_{av} = p_s \left(1 - \frac{a_2}{a_1}\right) = p_s \frac{x_{01} - x_{02} + 2\tau_l y}{x_{01} + x + \tau_l y}. \quad (2)$$

Poniendo las ecuaciones (1) y (2) juntas, se completa un modelo de aleteo extendido que rige la dinámica de los labios en la siringe de los oscinos.

En este modelo, las características acústicas de las soluciones están determinadas por parámetros significativos fisiológicamente. Suponiendo que los coeficientes en el término de la restitución del sistema (1) son proporcionales a la tensión de los músculos siríngeos ventrales (vS), este modelo es capaz de producir canto sintético realista. En [SAGM10], la actividad electromiográfica (EMG) registrada mediante electrodos implantados en los músculos vS se utilizaron para obtener una parametrización en función del tiempo de la restitución (la parte lineal de  $k(x)$ ). Conjuntamente, los datos de presión de aire del saco aéreo se utilizaron para obtener una parametrización de  $p_s$ . De esta manera pudo sintetizarse canto al conducir el modelo con datos fisiológicos reales.

**2.1. Separación fuente-filtro.** En el modelo descrito por el sistema de ecuaciones (1) y (2), se asume que la presión a la entrada del tracto es despreciable frente a la presión sub-siríngea ( $p_i = 0$ ). Es decir, si se cumple la *hipótesis de separación fuente-filtro*, la dinámica de las oscilaciones en la fuente es independiente de la del tracto.

Bajo ciertas condiciones, esta hipótesis no puede darse por cierta, y la presión a la entrada del tracto se introduce en la dinámica de la fuente mediante la presión inter-labial:

$$p_{av} = p_s + (p_s - p_i) \left(\frac{a_2}{a_1}\right) \quad (3)$$

Cuando esto ocurre, la dinámica del sistema fuente-tracto incrementa su riqueza y entre las soluciones oscilatorias que se agregan, hay algunas que introducen características acústicas no triviales en las soluciones.

## Parte 2

# Interacciones en la periferia: acople fuente filtro



## Buscando complejidad en interacciones sencillas

### 1. Fuentes de complejidad

En la parte anterior se presentó el canto de las aves oscinas como un comportamiento motor de resultado complejo. Las vocalizaciones de las distintas especies contienen una diversidad de características complejas, tanto en su estructura temporal como en su composición espectral [ZM04].

Esta tesis explora la hipótesis de que mucha de la complejidad se origina en la interacción entre el generador de patrones motores en el sistema nervioso central y el sistema complejo que es la periferia biomecánica no lineal. Una fuente de complejidad está en la versatilidad de las posibilidades de control motor de la siringe bipartita: la siringe de los oscinos está compuesta por dos válvulas labiales que pueden ser controladas de modo independiente. Al explotar mediante un programa motor adecuado esta propiedad de la fuente, se pueden producir sonidos complejos con la suma de las contribuciones de cada lado. Más allá de las características complejas que pueden obtenerse mediante un programa motor no trivial comandando las dos fuentes, ciertas propiedades biomecánicas intrínsecas de las estructuras vibrantes que producen sonido pueden contribuir a una riqueza de efectos en la dinámica que haga mella en las vocalizaciones aún a partir de gestos motores simples [ZRS08].

Una fuente de dinámica compleja en la periferia son las interacciones acústicas entre la fuente de sonido y el tracto vocal [ZRS08, AM09]. En este capítulo comenzamos a explorar sistemáticamente esta hipótesis, proponiendo un modelo matemático que de cuenta de la física del sistema fuente de sonido-filtro, y abra camino a la determinación de mecanismos dinámicos que originen complejidad relacionada a propiedades acústicas no triviales en vocalizaciones.

Introducimos un modelo sencillo para la fuente, que interactúa acústicamente con un tubo que hace las veces de tráquea. Proponemos un modelo minimalista para la fuente, en el sentido de que su dinámica es de baja dimensión y presenta apenas una bifurcación en la que se crean soluciones oscilatorias. Observamos, mediante una inspección numérica, que la introducción de la interacción fuente-tracto es responsable de al menos una bifurcación de las soluciones del sistema que conduce a la coexistencia de soluciones periódicas. Esta bifurcación tiene la potencialidad de generar saltos de frecuencia en las vocalizaciones, requiriendo apenas un cambio suave en la longitud del tracto vocal.



---

## 2. Observaciones experimentales sugestivas

Tomando la pregunta de cuáles fenómenos no lineales en la biomecánica de la periferia pueden originar qué complejidades en las vocalizaciones, Zollinger *et. al.* tutoraron ruiseñores (*Mimus polyglottos* con cantos sintéticos que contenían una serie de propiedades temporales y espectrales no triviales [ZRS08]). El ruiseñor es un imitador que puede copiar cantos heteroespecíficos. Se presume que puede incluir a su repertorio cantos de otras especies mediante un proceso de aprendizaje motor a prueba y error, guiado por la realimentación auditiva [ZS04].

Al estudiar los mecanismos motores utilizados por el ruiseñor para producir estas vocalizaciones, identificaron contribuciones a la diversidad vocal debidas a fenómenos no lineales y aquellas generadas mediante el control activo de la siringe bipartita. Para ello registraron el canto junto con la presión en los sacos aéreos y el flujo de aire a través de cada lado de la siringe. En ese trabajo determinaron que es posible la ocurrencia unilateral de fenómenos como *saltos de frecuencia, subarmonicidad y caos determinista*.

Un par de ejemplos de saltos de frecuencia producidos unilateralmente se muestran en la figura 1. Ocurre una variación discontinua en la frecuencia de la vocalización, que no está asociada a una alternancia entre las dos fuentes de sonido en la siringe, lo que pudo determinarse al observar el flujo en cada uno de los lados de la siringe durante la fonación [ZRS08]. De este modo se identifica que el origen de saltos en frecuencia y otras características complejas de las vocalizaciones (como subarmonicidad y caos determinista) no está en la alternancia entre ambas fuentes, ni correlacionado con alteraciones rápidas en el flujo aéreo o la presión del saco aéreo.

En las conclusiones del trabajo se especula que procesos biomecánicos pasivos, como la ocurrencia de fenómenos no lineales en una fuente, pueden derivar en la producción de estas vocalizaciones.

## 3. Un modelo no tan simple de acople fuente-tracto

Entre la fuente sonora y el medio se encuentra el tracto. La presión inter-valvular <sup>1</sup>, que proporciona la fuerza que motoriza las oscilaciones de los labios, depende de la presión en la entrada del tracto. De esta manera, el tracto es capaz de afectar el movimiento labial. En los seres humanos, se ha observado que la dinámica de las cuerdas vocales es independiente del tracto (salvo en algunas situaciones excepcionales [TRP08, Tit08]). Más allá de la hipótesis de separación fuente de sonido-filtro, sin embargo, la consideración de las interacciones entre la fuente y el filtro añade una gran complejidad a la periferia biomecánica responsable de la generación de sonido [LM05, AM09].

Uno de los modelos más utilizados para representar la transferencia de energía de un flujo de aire al tejido que oscila se conoce como el modelo de dos masas. Fue introducido por primera vez por Ishizaka y Flanagan [IF73] para dar cuenta de la generación de sonidos que componen la voz en humanos. Se estudió detalladamente por un número de investigadores [HBTS95, PHvH<sup>+</sup>94] y se aplicó al modelado matemático del órgano fonador

---

<sup>1</sup>La presión a través de la válvula que forman los labios: inter-glotal en las cuerdas vocales humanas, inter-siríngea en los oscinos.

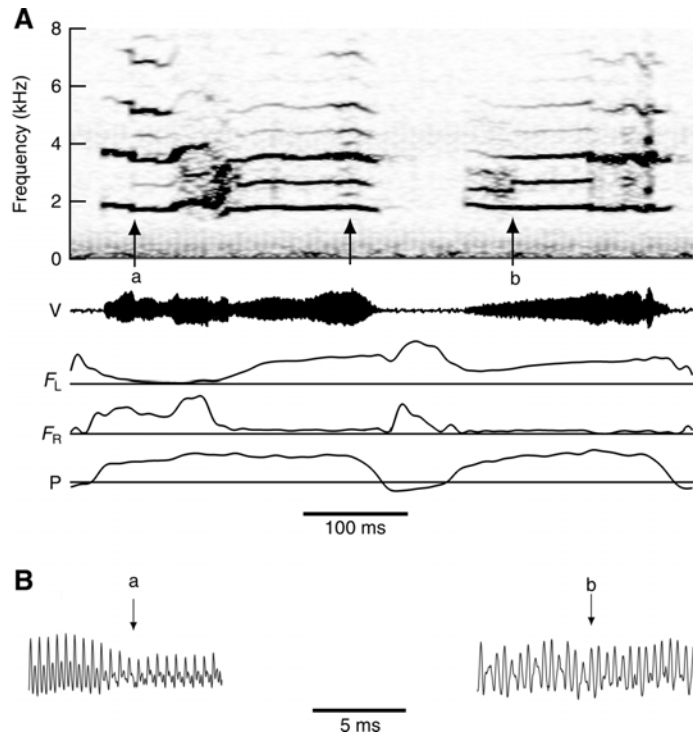


FIGURA 1. Ejemplos de saltos de frecuencia en el canto del ruiseñor. (A) Ocurren saltos de frecuencia con flujo de aire solo en el lado derecho (flecha a) o izquierdo (flecha b) de la siringe. (B) Vista expandida de la onda de sonido en los instantes que rodean los saltos de frecuencia, mostrando los cambios repentinos en los puntos de bifurcación.  $F_L$  y  $F_R$  es la tasa de flujo aéreo a través de los lados izquierdo y derecho de la siringe, respectivamente;  $V$  representa la amplitud de la onda de sonido;  $P$  la presión en el saco aéreo torácico. Adaptado de [ZRS08]

humano, así como al de las aves. Con diversas variaciones y adaptaciones, asume básicamente que el tejido oscilante en la fuente de sonido puede ser descrito en términos de dos masas, unidas entre sí por un resorte.

Este modelo de la fuente de sonido se ha utilizado por Hatzikirou *et. al.* para inspeccionar la dinámica de una fuente de sonido, cuando está acoplada a un tracto [HFH06]. Mediante exploraciones numéricas, en ese trabajo se muestra que para un sistema fuente-tracto modelado de esta forma, es de esperarse un escenario dinámico rico para las soluciones oscilatorias. Entre otros, encuentran saltos de frecuencia, desdoblamiento de periodo y caos determinista.

En este tipo de modelo es difícil, sin embargo, discernir si la complejidad del comportamiento es originada por la fuente o la interacción con el tracto. Dado que se trata de un modelo de cuatro dimensiones para la fuente (dos dimensiones para cada una de las masas), puede ocurrir dinámica compleja

---

incluso cuando la interacción con el tracto se desprecia. Dificultades adicionales surgen, en este enfoque, cuando se intenta realizar cálculos analíticos para conocer los orígenes de la dinámica de los fenómenos no lineales observados en las exploraciones numéricas.

#### 4. La inspiración en los oscinos: modelo de fuente sencilla acoplada al tracto

La dinámica de la fuente de sonido de los oscinos, no lineal en su naturaleza, exhibe fenómenos complejos que pueden introducir complejidad en las vocalizaciones incluso cuando se conduce con instrucciones fisiológicas simples.

Entre las diversas conclusiones de los trabajos de Zollinger *et al.* [ZRS08], se destaca la ocurrencia unilateral de fenómenos no lineales, tales como los saltos de frecuencia. Se observó también que estos hechos no son consistentes con las fluctuaciones de la presión del saco aéreo, ni con el flujo de aire a través de siringe, en apoyo a la hipótesis de que su ocurrencia no requiere gestos motores complejos.

En un modelo de órgano vocal capaz de dar cuenta de la generación de sonidos complejos, es difícil separar la complejidad asociada con el hecho de que los labios pueden presentar modos oscilatorios transporte complejos, de los fenómenos estrictamente asociadas con el acoplamiento acústico entre la fuente de sonido y el tracto [HFH06].

Con el objetivo de determinar si puede ocurrir comportamiento complejo debido con exclusividad a la interacción fuente-tracto, presentamos un modelo minimal en el que la fuente, cuando está desacoplada del tracto, sólo puede atravesar una bifurcación de Hopf [AM09]. Este modelo de fuente se inspira en el modelo de aleteo extendido con el que modelamos la siringe de los oscinos. El mecanismo físico de producción de sonido propuesto en este modelo es el mismo que el que se propone para explicar las oscilaciones en las cuerdas vocales humanas [Tit88, HBTS95, Tit00]. De esta manera, los fenómenos no lineales adicionales que tienen lugar cuando se introduce el acoplamiento en el modelo pueden ser identificado como una consecuencia de la interacción fuente-tracto.

En particular, al llevar el sistema a un límite altamente disipativo y al proponer el estudio de las ecuaciones de fase del sistema, se pueden encontrar las expresiones analíticas para la periodicidad de las soluciones [AM09]. Se pueden proponer condiciones sobre los parámetros que describen la longitud del tracto para la coexistencia de soluciones periódicas, y proponer un mecanismo para dar cuenta de la ocurrencia de saltos en la frecuencia en las vocalizaciones.

Tomando otro camino y aproximando al sistema que describe la dinámica de la fuente por un sistema débilmente no lineal (*weakly nonlinear*), se podrán obtener expresiones analíticas para la amplitud del sonido [APM11]. Con estas expresiones podremos explorar sistemáticamente los efectos introducidos en ella por el acoplamiento.

**4.1. Modelo de la fuente.** Teniendo esto en cuenta, proponemos aquí una versión minimalista del modelo de aleteo extendido. Este modelo tiene la ventaja de que fenómenos no lineales asociados con el acoplamiento

---

se identifican fácilmente. Luego, por medio de aproximaciones adecuadas, estaremos en condiciones de explorar sistemáticamente los efectos introducidos por el acople en la dinámica de la frecuencia y de la amplitud de los sonidos producidos por el sistema.

El modelo minimalista para la fuente se construye sobre la base del modelo de aleteo extendido presentado en [LGM02], revisado en la sección 2 del capítulo 2. Se rige por una ecuación de segundo orden [AM09]: la segunda ley de Newton para el apartamiento del equilibrio del centro de masa de un labio  $x$ :

$$\begin{cases} \dot{x} &= y \\ \dot{y} &= -kx - \beta y - cx^2y + p_i + (p_s - p_i)f(x, y), \end{cases}$$

donde, en la segunda ecuación, el primer término describe la restitución elástica del labio, el segundo término representa la disipación, y el tercero una saturación no lineal que limita el movimiento labial [AM09, APM11].

Notar que en esta versión del modelo, escribimos más genéricamente la presión inter-labial como una función de las presiones sub y supra-siríngicas ( $p_s$  y  $p_i$ ) y una función genérica  $f(x, y)$  que describe el perfil de la válvula siríngica en términos de la posición y la velocidad del punto medio del labio.

El sistema es forzado por los dos últimos términos. Ellos representan la presión promedio a través de la válvula siríngica, escrita en términos de la presión sub-siríngica  $p_s$ , y la presión en la entrada del tracto  $p_i$  (todas las presiones se definen por unidad de masa por unidad de área del labio). En el término de forzado,  $f(x, y)$  es una función de la geometría de las membranas que depende de la relación entre las secciones de los extremos traqueal y bronquial de la válvula labial. La diferencia de fase experimentalmente observada entre la parte superior e inferior de los labios se introduce en esta función [Tit88, LGM02, LM05, AM09].

Una forma equivalente de plantear el requisito para que ocurra el movimiento de “aleteo” es que el promedio de la presión entre los labios sea más cercano a la presión bronquial cuando los labios presentan un perfil convergente, y más cercanos a la presión atmosférica cuando son divergentes. Por lo tanto, la fuerza va en la misma dirección que la velocidad de desplazamiento de los labios, y es capaz de superar la disipación para valores suficientemente altos de la presión sub-siríngica. Este requisito se cumple si  $f(x, y)$  es proporcional a la la velocidad de los labios (*i. e.*,  $f(x, y) = y/v_{char}$ , con  $v_{char}$  una velocidad característica) [LGM02].

En contraste con los modelos de masas, más detallados, nuestro sistema restringe la dinámica de la fuente a un modo espacial simple. A pesar de su simplicidad, este modelo alcanza a capturar la complejidad del fenómeno de modo suficiente como para reproducir comportamiento vocal realista. En un trabajo anterior, se exploró su dinámica en el espacio de parámetros  $(p_s, k)$ , y se encontró que es capaz de dar cuenta de los mecanismos de producción de sonido del cardenal norte (*Cardinalis cardinalis*) [MGGS03].

Con la  $f(x, y) = y/v_{char}$  propuesta, el sistema presenta un punto fijo en  $(x, y) = (0, 0)$ . Luego de un cambio de escala ( $t \rightarrow t/\gamma$ ,  $y \rightarrow \gamma y$ ), y fijando

$v_{char} = 1$  por simplicidad, escribimos

$$\begin{cases} \dot{x} &= y \\ \dot{y} &= -k\gamma^2 x + \gamma(p_s - \beta)y - \gamma c x^2 y + \gamma p_i(\gamma - y). \end{cases} \quad (4)$$

Asumir que se cumple la *hipótesis de separación fuente-filtro* significa que la presión en la base del tracto se considera despreciable en el término de forzado del sistema. Esto se expresa mediante el establecimiento de  $p_i = 0$  en (4). Para algunos valores de los parámetros  $(p_s, \beta)$ , la fuerza ejercida por la presión supera la disipación y se produce una bifurcación de Hopf: el punto fijo se vuelve inestable y un ciclo límite nace con amplitud cero y finito de frecuencias [Str00]. Más allá de la bifurcación, el punto medio de los labios comienza a oscilar en torno a su posición de equilibrio.

## 5. Un tracto simple y acoplado

El acople entre la fuente y el tracto es introducido al considerar  $p_i \neq 0$ . Cuando la dinámica de  $p_i = p_i(x, y, t)$  se introduce, se agrega complejidad a las ecuaciones de movimiento de los labios. Si la válvula labial está acoplada a un tubo, la presión en la entrada del tracto  $p_i$  se verá afectada por las reflexiones de la onda de sonido a la salida. Hay una contribución al término de forzado que proviene de una realimentación y depende del valor de  $p_i$  en un tiempo anterior, que es función de la longitud  $L$  del tracto.

Derivamos una forma funcional de  $p_i$  que tiene en cuenta esa realimentación [AM09]. Asumimos dos contribuciones principales a la presión supra-siríngica  $p_i$ : una debida a las fluctuaciones originadas en la siringe e inyectadas en el tubo, y otra debida a la realimentación.

Para fluctuaciones del flujo del orden de los  $kHz$  y sección y el tubo del orden de los  $mm$ , la contribución de las fluctuaciones del flujo siríngico  $U_s$  se pueden escribir como  $p_+ = \rho_0 v_s U_s / A_i$ , donde  $\rho_0$  es la densidad del aire no perturbado,  $A_i$  la sección de la entrada del tracto, y  $v_s$  la velocidad del sonido. La velocidad media del aire en la siringe viene dada por una versión fenomenológicamente corregida de la ley Bernoulli,  $V_m = \sqrt{\frac{2}{p_s k_t \rho_0}}$ , en la que  $k_t$  representa el coeficiente de presión trans-siríngica (análogo a la presión trans-glotal descrita por Titze en la derivación del modelo de aleteo para oscilaciones en las cuerdas vocales de humanos [Tit88]). El flujo siríngico se puede aproximar como  $U_s = V_m A_m$ , donde el área media de la válvula siríngica  $A_m$  es proporcional al desplazamiento del equilibrio del punto medio de los labios  $x$ . Por lo tanto, se puede escribir la contribución a  $p_i$  por fluctuaciones en el flujo siríngico como  $p_+ = \alpha \sqrt{p_s} x$ , donde  $\alpha$  es inversamente proporcional a la superficie del tubo.

La otra contribución a la presión en la entrada del tracto proviene de la reflexión a la salida. Modelamos el tracto como un tubo uniforme de longitud  $L$ , abierto en el extremo más alejado de la fuente. Una onda de sonido entrante al tubo por el extremo que linda con la fuente será transmitida parcialmente en el otro extremo, y en parte será reflejada con un coeficiente  $r$ . La reflexión de una onda de sonido que ingresa al tubo en un momento dado contribuirá a la presión en la entrada del mismo con un retraso de  $\tau = 2L/v_s$  (el tiempo que tarda la onda en propagarse al otro extremo y de

---

vuelta), y signo opuesto. En esa expresión,  $v_s$  es la velocidad del sonido en el medio.

Teniendo en cuenta ambas contribuciones, la presión supra-siríngea puede ser escrita como

$$p_i(t) = \alpha\sqrt{p_s}x - rp_i(t - \tau). \quad (5)$$

Una exploración numérica del sistema con realimentación (5) muestra una variedad de soluciones dinámicas. En particular, es posible la coexistencia de soluciones periódicas.

## 6. La complejidad de la dimensión infinita: saltos de frecuencia

El sistema tracto-fuente acoplados descrito por las ecuaciones (4, 5) es un sistema de ecuaciones diferenciales con retardo (DDE, por las siglas en inglés *Delay Differential Equation*). Es un sistema de dimensión infinita: la condición inicial para obtener una trayectoria en el espacio de fases comenzando en un  $t_0$  incluye al conjunto infinito de valores  $p_i(t_0 - t')$ , con  $t' \in [0, \tau]$  [Gra87]. Realizamos exploraciones numéricas para encontrar fenómenos no lineales en este sistema. Se ha reportado en trabajos recientes que el canto es acompañado por un gesto motor responsable de la adaptación de las dimensiones del tracto vocal, coordinado con la frecuencia fundamental generada en la siringe [RSFB06, SM02, ZM04]. Por esta razón, hemos restringido nuestras exploraciones numéricas del modelo a los parámetros que describen el tubo de manera que  $\tau = \frac{2L}{v_s}$  sea del mismo orden de magnitud que el periodo  $T_0$  de las soluciones de la fuente desacoplada del tracto. Una primera inspección de las soluciones se realizó por medio de integración numérica del modelo para  $\tau \in [0, \frac{5}{2}T_0]$ .

Integramos el sistema con realimentación retardada (4, 5) para valores de  $\tau$  dentro de un rango, con diferentes condiciones iniciales y, después de un transitorio, intersectamos las trayectorias en el espacio de fases  $(x, y)$  con una sección de Poincaré. Los puntos de intersección  $x_p$  de las trayectorias con la sección de Poincaré se grafican en función de  $\tau$  en la figura 2a. Una situación particularmente interesante se produce cuando  $\tau$  toma valores cercanos al periodo del ciclo límite no perturbado: la coexistencia de soluciones periódicas. Esta histéresis entre los ciclos límite se muestra en la figura 2b, que es una ampliación del tramo entre las flechas de la figura 2a. En la figura 2c se ilustran los periodos de las soluciones numéricas en la región detallada.

La coexistencia de soluciones periódicas es una de las principales características exhibidas por los osciladores de ciclo límite sujetos a realimentación retardada [Gra87, EG08, BBM<sup>+</sup>03]. Además, esta ocurre cercana, en el espacio de parámetros, a la zona donde la frecuencia fundamental de la fuente coincide con la resonancia del tracto. Esta observación concuerda con la coordinación reportada entre la frecuencia fundamental y la geometría del tracto. La hipótesis que exploramos a continuación mediante el estudio analítico de la dinámica de este modelo, es que ciertos efectos dinámicos asociados a esta coexistencia hacen posible la ocurrencia de vocalizaciones complejas requiriendo un gesto motor suave.

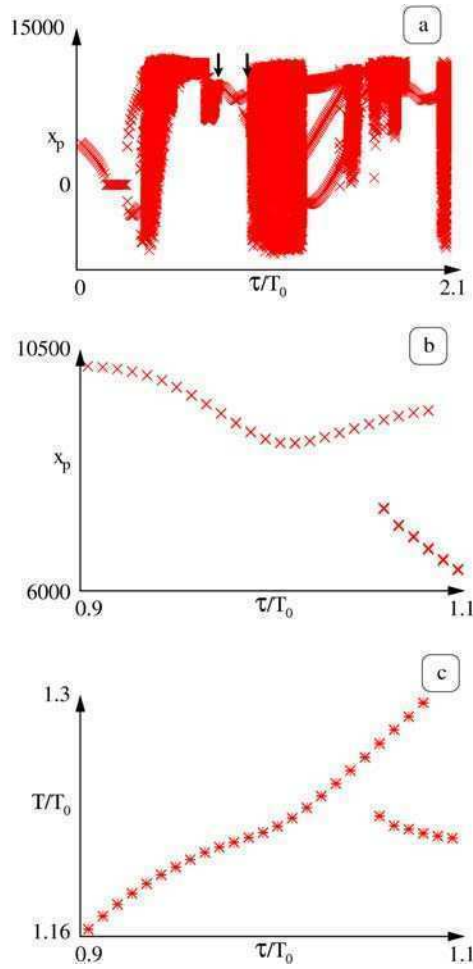


FIGURA 2. Valores de  $x(t)$  en la intersección de trayectorias en el espacio de fases  $(x, y)$  del sistema (4) con realimentación (5) y la línea  $y = 5000x$  (a). Una mirada detallada de la sección entre las flechas sugiere coexistencia de soluciones de distinto periodo (b). Al medir los periodos  $T$  de trayectorias en la región detallada, para distintas condiciones iniciales, corroboramos la coexistencia (c). Para los parámetros usados en la integración numérica,  $(\gamma, \beta, p_s, k) = (7000\pi, 1, 0, 0, 3, 1, 0)$ , el ciclo límite encontrado cuando  $p_i = 0$  (fuente y filtro separados) tiene periodo  $T_0 = 1/3500s$ . Para la búsqueda de coexistencia de soluciones periódicas se fijó  $p_s = 0,3$  y  $(\alpha, r) = (1, 0, 0, 4)$ .

## La dinámica de la frecuencia: disipación infinita

En la exploración numérica del modelo de aleteo extendido para la fuente acoplada al tracto encontramos un escenario dinámico rico, que incluye una región de coexistencia de soluciones periódicas. Este fenómeno, identificado con la interacción con el filtro, ocurre cuando la frecuencia del sonido producido por la fuente es cercana a la resonancia del tubo con el que modelamos la tráquea. Por otro lado, hay evidencia de que algunas especies coordinan la geometría del tracto con la frecuencia fundamental de la vocalización [RSFB06, SM02, ZM04]. Por estos tres motivos, en este capítulo nos dedicamos a desentrañar los orígenes dinámicos de esa coexistencia, y su relación con los saltos de frecuencia.

Ya que el modelo de la fuente es de baja dimensión (dimensión 2), es relativamente fácil reducir el sistema fuente-filtro a una forma que permite estudiar la dinámica de sus soluciones periódicas analíticamente [Gra87]. Para esto, aproximaremos el sistema que describe a la fuente por un oscilador con disipación infinita.

En este límite, las perturbaciones introducidas por el término con retardo son pequeñas perturbaciones a las trayectorias estables del oscilador libre [Str00, EG08]. Cuando la disipación es alta, además, los apartamientos de las trayectorias respecto de las soluciones oscilatorias estables decaen rápidamente hacia el ciclo límite, y tiene sentido describir las oscilaciones solamente en términos de una variable de fase [PRK01, EG08]. Estudiando analíticamente el efecto de la perturbación introducida por la realimentación retardada en esa variable de fase, podremos determinar condiciones para la coexistencia de soluciones periódicas estables en el sistema fuente-tracto.

### 1. La aproximación

Antes de introducir el acople entre la fuente y el tracto, la dinámica de los labios es descrita por

$$\begin{cases} \dot{x} &= y, \\ \dot{y} &= -k\gamma^2 x + \gamma(p_s - \beta)y - \gamma x^2 y, \end{cases} \quad (6)$$

que es la forma estándar de

$$\begin{cases} \dot{u} &= v - F(u) = v - \gamma u^3/3 + \gamma(p_s - \beta)u, \\ \dot{v} &= -k\gamma^2 u. \end{cases} \quad (7)$$

Es decir, estos dos sistemas de ecuaciones son equivalentes si  $x = u$  y  $v = \dot{x} - \gamma(p_s - \beta)u = 0$  [LGM02]. El sistema de ecuaciones (7) describe la dinámica del oscilador de Van der Pol, un modelo paradigmático de oscilaciones de relajación. Su dinámica está regida por la nulclina cúbica  $v - F(u) = v - \gamma u^3/3 + \gamma(p_s - \beta)u = 0$ . Cualquier trayectoria en el espacio de fases  $(u, v)$



se aproxima rápidamente hacia la nulclina, y la recorre lentamente hasta alcanzar un extremo. Luego de alcanzar ese punto, salta hacia la otra rama de la nulclina. Esto es seguido de otro recorrido lento hasta alcanzar el extremo de esa otra rama de la nulclina, donde otro salto ocurre y la dinámica se sucede periódicamente [Str00].

La dinámica es equivalente si la nulclina cúbica se aproxima con segmentos rectos:

$$f(u) = \begin{cases} u + 2\sqrt{p_s - \beta} & x < -\sqrt{p_s - \beta}, \\ -u & |u| \leq \sqrt{p_s - \beta}, \\ u - 2\sqrt{p_s - \beta} & x > \sqrt{p_s - \beta}. \end{cases}$$

El oscilador de relajación que presenta esta nulclina y la misma escala temporal que nuestro modelo sigue el sistema de ecuaciones

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -k\gamma^2 x + \gamma Sg((p_s - \beta) - x^2))y, \end{cases} \quad (8)$$

donde  $Sg(z)$  significa *signo de z*:  $Sg(z) = 1$  si  $z \geq 0$ ,  $Sg(z) = -1$  de otro modo.

El uso de este modelo en vez de (6) nos permite obtener expresiones analíticas para el movimiento a lo largo de la nulclina, lo que será una ventaja a la hora de buscar condiciones analíticas para el fenómeno en el que estamos interesados. Aún así, es pertinente notar que una disipación que se enciende cada vez que el apartamiento de  $x$  del equilibrio excede un cierto umbral (como es el caso en el modelo descrito por la ecuación (8)), posiblemente es incluso más realista que la disipación no lineal representada por el término  $x^2 y$  en (6) para modelar las colisiones que confinan a los labios a válvula siríngica.

Por estas razones, nos concentraremos en el estudio de la dinámica de una fuente de sonido, cuya dinámica viene descrita por el sistema (8), acoplada a un tracto sencillo (con  $r < 1$ ), en el sistema resultante gobernado por

$$\ddot{x} + \nu Sg(x^2 - (p_s - \beta)) \dot{x} + x = \delta(1 - \dot{x}/v_{char})(x - rx(t - \nu\tau)), \quad (9)$$

donde  $\gamma^2 = 1/k$ , y  $\nu = 1/\gamma$ .

## 2. Dinámica de los ciclos límite

La ventaja de trabajar con el sistema altamente disipativo descrito por la ecuación (9) es que el oscilador libre admite una solución sencilla. La evolución a lo largo de las ramas rectas de la nulclina puede expresarse analíticamente, y por lo tanto el periodo de sus soluciones oscilatorias puede aproximarse fácilmente:  $T_0 \simeq 2\nu \ln 3$ . El parámetro  $\nu$  determina la escala temporal del retardo  $\tau$ , ya que queremos estudiar cómo afecta al periodo la realimentación con retardos del mismo orden de magnitud. Una técnica empleada en [Gra87] puede ser adaptada para derivar una ecuación de fase con retardo (ecuación DDE de fase), que será analizada en términos del retardo  $\tau$  en el límite de alta disipación y realimentación débil.

Como se mencionó en la sección anterior, la ecuación (9) describe un oscilador de relajación lineal a trozos, sujeto a realimentación retardada. La naturaleza de sus soluciones periódicas puede ser inspeccionada analíticamente. Empezaremos por estudiar las soluciones no perturbadas ( $\delta = 0$ ).

Luego, definiremos una variable de fase se para el oscilador libre y adaptaremos un método que se presenta en [EG08] para obtener una ecuación de fase perturbada para el sistema cuando se introduce la realimentación. De esta ecuación fase, se pueden sacar conclusiones sobre cómo se ven afectadas las trayectorias en el ciclo límite por la introducción de la realimentación.

En la representación de Liénard, el sistema representado por (9) toma la forma de un sistema de dos ecuaciones diferenciales de primer orden. Cambiando la escala de tiempo como  $t \rightarrow t/\nu$  y eligiendo  $p_s - \beta = 1$  como parámetros consistentes con soluciones oscilatorias, llegamos a

$$\begin{cases} \epsilon u' &= v - f(u), \\ v' &= -u + \delta(1 - \nu(v - f(u)))(u - ru(t - \tau)). \end{cases} \quad (10)$$

donde  $\epsilon = 1/\nu^2$  y  $f(u)$  es una función lineal a trozos que resulta en  $Sg(u^2 - 1)$  al ser derivada respecto de  $u$ :

$$f(u) = \begin{cases} u + 2 & u < -1, \\ -u & |u| \leq 1, \\ u - 2 & u > 1. \end{cases}$$

Para obtener una definición de la fase del oscilador libre, es preciso prestar atención a su ciclo límite. Fijamos  $\delta = 0$  en (10) y encontramos que a disipación infinita ( $\epsilon \rightarrow 0$ ), las oscilaciones se aproximan a un límite discontinuo en el que las trayectorias en el espacio de fases  $(u, v)$  satisfacen

$$u_0(t) = \begin{cases} 3e^{-t} & 0 < t \leq t_0, \\ -3e^{-(t-t_0)} & t_0 < t \leq T_0, \end{cases} \quad v_0(t) = \begin{cases} u_0 - 2 & 0 < t \leq t_0, \\ u_0 + 2 & t_0 < t \leq T_0, \end{cases} \quad (11)$$

Aquí,  $t_0 = \ln 3$  y  $T_0 = 2t_0$  es el periodo de las oscilaciones libres. Las ecuaciones (11) describen trayectorias a lo largo de dos secciones del ciclo límite de (10) cuando la disipación es alta. Cuando  $u > 1$ , la solución decrece monótonamente a lo largo del ciclo hasta el punto  $(1, -1)$ , donde salta a la rama izquierda, cayendo en  $(-3, -1)$ . Una vez allí, crece monótonamente y se desprende de esa rama en  $(-1, 1)$ , alcanzando la rama derecha en  $(3, 1)$  y recomenzando el ciclo. Se puede definir una fase que aumente monótonamente para el oscilador libre. De este modo, cuando se enciende la perturbación (la realimentación), y dado que la amplitud de las órbitas periódicas permanecerá casi inalterada, la perturbación en el movimiento será reflejada por la dinámica de la fase [PRK01]. Despreciando el tiempo que toma a las trayectorias el salto de una rama del ciclo a la otra, la fase buscada crecerá uniformemente mientras satisfaga la ecuación

$$\frac{du_0}{d\phi} = -u_0. \quad (12)$$

Las soluciones del sistema con realimentación en el límite de disipación infinita ( $\epsilon \rightarrow 0$ ) consisten de una parametrización diferente en el tiempo del ciclo límite no perturbado, *i.e.*,

$$u = u_0[\phi(t)]. \quad (13)$$

Cuando se introduce la realimentación fijando  $0 < \delta \ll 1$ , se encuentra una ecuación que da cuenta de la dinámica de la fase. En una primera aproximación, la amplitud de la órbita no depende de  $\delta$  pero pueden ocurrir

cambios en la fase. Para este sistema con disipación infinita, la forma de Liénard con realimentación, análoga a (10), toma la forma

$$\begin{cases} \dot{0} &= v - f(u) \\ \dot{v} &= -u + \delta(1 - \nu(v - f(u)))(u - ru(t - \tau)). \end{cases}$$

Por eliminación puede ser reducida a

$$\dot{u} = -u + \delta(u - u(t - \tau)), \quad (14)$$

junto con las condiciones sobre los saltos entre las ramas:  $u = -3$  si  $u < 1$  y  $u = 3$  si  $u > -1$ . Aplicando la definición de fase (13) y la condición (12), la DDE (14) puede ser escrita como

$$\frac{d\phi}{dt} = 1 - \delta + \delta \frac{u_0[\phi(t - \tau)]}{u_0[\phi(t)]}.$$

En la búsqueda de una solución a esta ecuación de fase, usamos un procedimiento perturbativo que tiene en cuenta el hecho de que hay dos escalas de tiempo en el problema [Str00]. Una escala rápida  $t$  es adecuada para describir las trayectorias sobre el ciclo límite del oscilador libre. Dada la naturaleza altamente disipativa del sistema, una escala temporal mucho más lenta  $s = \delta t$  es requerida para describir las perturbaciones de las trayectorias introducidas por la realimentación. En el *método de dos escalas temporales*, estas escalas son tratadas como variables independientes [Str00]. Así, proponemos una solución perturbativa  $\phi(t, \delta) = \phi_0(t, s) + \delta\phi_1(t, s) + O(\delta^2)$ . La derivación respecto del tiempo, siguiendo la regla de la cadena, deviene  $\frac{d}{dt} \rightarrow \frac{d}{dt} + \delta \frac{d}{ds}$ .

La ecuación a orden cero en  $\delta$  queda  $\frac{d\phi_0}{dt} = 1$ , y su solución es

$$\phi_0(t, s) = 1 + \psi(s).$$

Usando este resultado para continuar a orden  $\delta$ , llegamos a la ecuación

$$\frac{d\phi_1}{dt} = -1 - \frac{d\psi}{ds} + \frac{u_0[t - \tau + \psi(s - \delta\tau)]}{u_0[t + \psi(s)]}.$$

Como  $\phi_1(t, s)$  es la corrección a primer orden de la fase perturbada por la realimentación, queremos que su promedio respecto del tiempo rápido  $t$  a lo largo de cada periodo de oscilación sea cero. De otro modo, la fase del sistema perturbado divergería de la del oscilador libre. Esto impone la *condición de solubilidad (solvability condition)*  $\frac{1}{T} \int_0^T \frac{d\phi_1}{dt} dt = 0$ , o

$$\frac{d\psi}{ds} = -1 + F(\Delta), \quad (15)$$

donde

$$F(\Delta) = \frac{1}{T} \int_0^T \frac{x(\zeta + \Delta)}{x(\zeta)} d\zeta,$$

y  $\Delta = \psi(s - \delta\tau) - \psi(s) - \tau$  permanece constante a lo largo de la integral sobre  $\zeta = t + \psi(s)$ . Si se propone una solución lineal para (15):

$$\psi(s) = \sigma s + \psi_0, \quad (16)$$

se encuentra

$$\Delta = -\tau(1 + \delta\sigma), \quad (17)$$

y  $F(\Delta)$  puede calcularse. En general  $F(\Delta)$  tomará una forma diferente, dependiendo de si  $u_0$  cae sobre la rama derecha o izquierda del ciclo límite. En este problema,

$$F(\Delta) = (-1)^n e^{-\Delta - nt_0} \left[ \left( \frac{4}{3}n + 1 \right) + \frac{4}{3}\Delta \right], \quad -(n+1)t_0 \leq \Delta < -nt_0,$$

con  $n = 0, 1, 2, \dots$ . Insertando el *ansatz* (16) en (15), encontramos una ecuación que relaciona los parámetros  $\sigma$  y  $\Delta$ :

$$\sigma = 1 - F(\Delta). \quad (18)$$

Las expresiones (17) y (18) conducen a

$$\tau = \frac{-\Delta}{1 - \delta(1 - F(\Delta))}. \quad (19)$$

El periodo de la órbita perturbada se calcula como

$$T = \frac{2t_0}{1 + \delta\sigma} = \frac{2t_0}{1 + \delta(1 - F(\Delta))}, \quad (20)$$

con  $\Delta$  tal que satisfaga (19). El modo de calcular el periodo de la solución dada una fortaleza de la realimentación  $\delta$  consiste en encontrar una solución para (19) (*i.e.*, los valores de  $(\tau, \Delta)$  que cumplirían la ecuación), y luego introducir  $\Delta$  en esta última ecuación. En la figura 1 se grafica una línea de soluciones de (19) para  $\delta$  fijo. Para algunos valores de  $\tau$ , hay dos valores de  $\Delta$  que satisfacen la condición (esto es, para algunos rangos de  $\Delta$  la relación  $\tau(\Delta)$  implícita en la ecuación (19) es no-inyectiva). En esos casos,  $T$  admite dos soluciones. Los periodos  $T$  de las soluciones perturbadas, calculadas

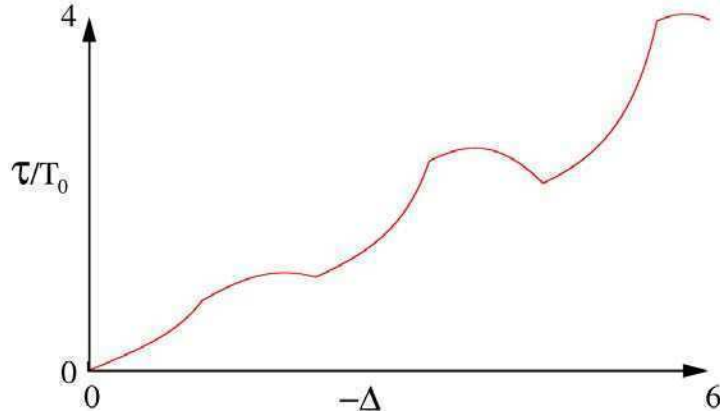


FIGURA 1. Soluciones de la ecuación (19) con  $\delta = 0,19$ . Cada par  $(\Delta, \tau)$  en la línea es luego utilizado para calcular el periodo de una solución perturbada usando la ecuación (20). De esta manera, soluciones periódicas coexisten cuando hay dos puntos en la línea con el mismo  $\tau$ .

como se describió, están representadas contra el retardo  $\tau$  con una línea sólida en la figura 2. Para ese valor de  $\delta$ , periodos de la solución numérica del sistema (10) son graficados como puntos en el mismo gráfico, a fin de ilustrar el acuerdo de los cálculos analíticos con la simulación numérica. (ver la leyenda de la figura 2 para los parámetros usados en la simulación). El

periodo exhibe biestabilidad para  $\tau$  suficientemente grandes. Cuando una región de biestabilidad es cruzada, ocurre un “salto” en  $T$ .

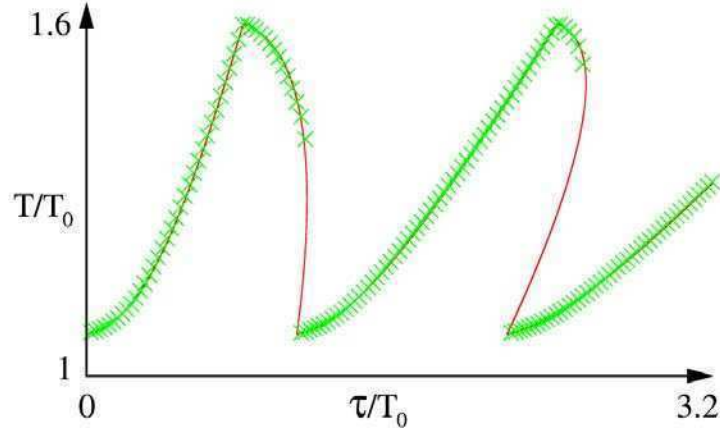


FIGURA 2. Período en función del retardo  $\tau$  para el sistema DDE (9) con  $\delta = 0,19$ . La línea es el gráfico del resultado analítico, obtenido mediante el método de dos escalas temporales en la ecuación de fase del sistema. Los puntos corresponden a medidas del periodo de simulaciones numéricas de (10) (en el límite  $\epsilon \rightarrow 0$ ; se fijó  $\nu = 100$ ).

### 3. Origen dinámico de los saltos de frecuencia

Los saltos en los periodos de las soluciones mencionadas en la sección anterior se asemejan a los saltos de frecuencia observados por Zollinger *et. al.* [ZRS08]. Esta característica no lineal de los sistemas con retroalimentación retardada es uno de los candidatos a ser responsables de la generación del fenómeno.

Mediante la sincronización de un gesto de presión adecuado con una excursión del parámetro  $\tau$  que incluye el cruce de una región de coexistencia, esperamos generar una sílaba sintética en el que se produce un salto de frecuencia. Un conjunto de ecuaciones similares a (4) ha sido previamente utilizado para generar sílabas sintéticas, al variar la presión  $p_s$  durante intervalos de tiempo, siguiendo diferentes patrones [ML05]. Aquí, vamos a presentar un patrón de presión muy simple, que consiste meramente en el aumento de la presión por encima del valor en el que los labios en la siringe comienzan a oscilar, y después de un intervalo de tiempo, el retorno a un valor por debajo del umbral en el que las membranas permanecen en un punto fijo estable (sin desplazamiento). Durante ese intervalo, la longitud del tracto  $L$  es incrementada linealmente, llevando al retraso  $\tau$  a cruzar una región de bi-estabilidad. Como consecuencia de ello, se producirá un salto en el periodo de las soluciones. Este efecto se refleja en las características acústicas de la sílaba sintética. Se produce, como se esperaba, un salto en la frecuencia del sonido generado, cuando la solución periódica correspondiente a la longitud de partida  $L_i$  ya no existe y las trayectorias en el espacio de fase son atraídas a la órbita periódica estable correspondiente a la longitud

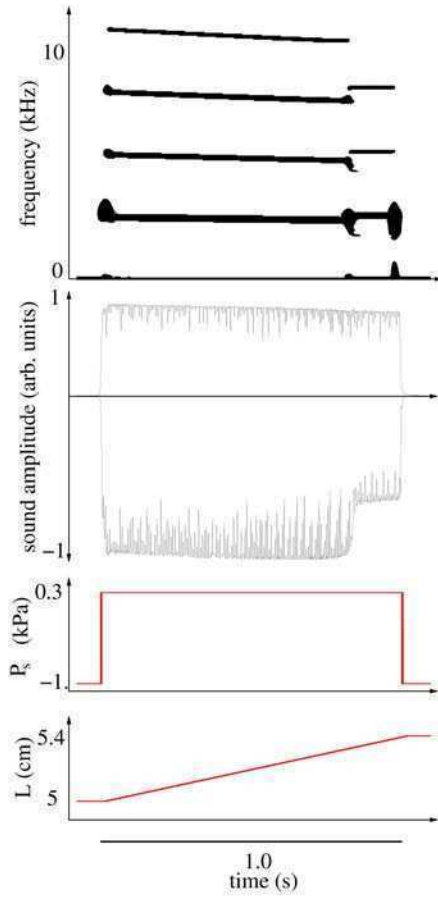


FIGURA 3. Sonido sintético producido integrando numéricamente las eqs. (4, 5). Sonido (segundo panel desde arriba) es producido cuando el sistema es conducido por un gesto de presión sencillo (tercer panel desde arriba). Al variar la longitud del tracto dentro de la región de coexistencia (panel inferior), el sonograma (panel superior) exhibe un salto de frecuencia. El sistema fue integrado con parámetros  $(\gamma, \beta, p_s, k) = (7000\pi, 1, 0, 0, 3, 1, 0)$  y coeficientes de la realimentación  $(\alpha, r) = (1, 0, 0, 4)$  para generar la sílaba ( $p_s = -1, 0$  corresponde a no fonación), y  $L$  se incrementó linealmente durante la fonación desde  $L_i = 5,01\text{cm}$  hasta  $L_f = 5,4\text{cm}$ , correspondientes a retardos  $\tau_i = 29,48\text{ms}$ ,  $\tau_f = 31,75\text{ms}$ .

final  $L_f$ . La figura 3 ilustra, de arriba a abajo, el sonograma del sonido sintetizado, la onda de sonido normalizada, el gesto de presión y el estiramiento del tracto vocal.

Hemos mostrado que, para longitudes tales que las resonancias del tracto son cercanas a la frecuencia natural de la fuente, se produce multi-estabilidad de soluciones periódicas. No todo salto en la frecuencia fundamental del canto de los pájaros se puede atribuir a este mecanismo. En algunos casos,

---

los saltos se corresponden con sonidos generados sucesivamente por cada una de las dos fuentes de sonido. Sin embargo, los sonidos con saltos de frecuencia generados de manera unilateral se han reportado recientemente [ZRS08]. El mecanismo que se ilustra en este capítulo muestra que en estos casos, los cambios acústicos rápidos no requieren necesariamente de control muscular rápido.

## La dinámica de la amplitud: la fuente *cuasi*-lineal

### 1. Motivación

**1.1. Antecedentes.** Con el objetivo de determinar si puede ocurrir comportamiento complejo debido solo a la interacción fuente-tracto, presentamos un modelo en el que la fuente, cuando está desacoplada del tracto, sólo puede atravesar una bifurcación de Hopf [AM09]. De esta manera, los fenómenos no lineales adicionales que tienen lugar cuando se introduce el acoplamiento en el modelo pueden ser identificados como una consecuencia de la interacción fuente-tracto.

Al llevar el sistema a un límite altamente disipativo y al proponernos el estudio de las ecuaciones de fase del sistema, pudimos encontrar las expresiones analíticas para la periodicidad de las soluciones. Se establecieron condiciones sobre los parámetros que describen la longitud del tracto para la coexistencia de soluciones periódicas, y se propuso un mecanismo para dar cuenta de la ocurrencia de saltos en la frecuencia de las vocalizaciones.

Los gestos motores que determinan las frecuencias fundamentales de las vocalizaciones de las aves oscinas son coordinadas con la geometría de varias partes del tracto vocal, tales como la longitud de la tráquea, el volumen de la cavidad esofágica oro-faríngea, o la apertura pico [RSFB06, GC04, AT10]. En muchos de los experimentos, sin embargo, esta actividad coordinada no da lugar a efectos no lineales tan evidentes como los saltos de frecuencia. Con el fin de determinar las contribuciones de la interacción fuente-tracto a la complejidad del canto, es útil obtener sus efectos sobre la amplitud del sonido, que es el observable más directo.

Teniendo esto en cuenta, aquí estudiamos el modelo presentado en [AM09] y en la sección 4 de una manera que nos permite observar los fenómenos no lineales en la amplitud del sonido. Este modelo tiene la ventaja de que los fenómenos no lineales asociados con el acoplamiento se diferencian fácilmente de los debidos a la complejidad de la fuente. Trabajando en el límite débilmente no lineal (*weakly nonlinear*), obtenemos expresiones analíticas para la amplitud del sonido. Con estas expresiones podemos explorar sistemáticamente los efectos introducidos en ella por el acoplamiento, prestando especial atención a las regiones donde la frecuencia del sonido producido en la fuente es cercana a las resonancias del tracto.

**1.2. Observaciones numéricas.** En este capítulo nos concentramos en la búsqueda de cambios cualitativos introducidos por el acople en la amplitud de los sonidos generados en el sistema. La presión acústica en la entrada del tracto es  $p_i(t)$ . Por lo tanto, la onda transmitida parcialmente en la salida del tracto en un momento dado  $t$  es  $p(t) = (1 - r)p_i(t - \tau/2)$ . Esta cantidad es la presión acústica en la salida del sistema. Para encontrar la



amplitud de un sonido sintetizado se calcula  $p(t)$  integrando numéricamente las ecuaciones (4, 5) y, después de un tiempo transitorio suficiente (*i.e.*, tiempo suficiente para que se observen oscilaciones de amplitud constante), encontrar el máximo de su norma. Esta cantidad,  $|p|_{max}$ , es en adelante llamada la “amplitud del sonido”.

Nos centramos en la región donde la frecuencia fundamental del sonido producido por la fuente sin perturbar  $f_0 = \omega_0/2\pi$  es cercana a la resonancia del tracto, que es donde encontramos previamente coexistencia ([AM09] y capítulo 4). Ya que nos concentramos en los cambios cualitativos introducidos por el acoplamiento fuente-tracto en la dinámica del sistema, establecemos valores adimensionales para los parámetros, que nos permitan detectar las bifurcaciones en las soluciones oscilatorias a través de la exploración numérica.

Nos encontramos con un comportamiento no trivial de la amplitud, dependiendo de la fortaleza del acople (coeficiente  $\alpha$ ). Estos resultados se resumen en las figuras 1 y 2. En la figura 1 representamos  $|p|_{max}^2$  vs.  $(k, \alpha)$ .

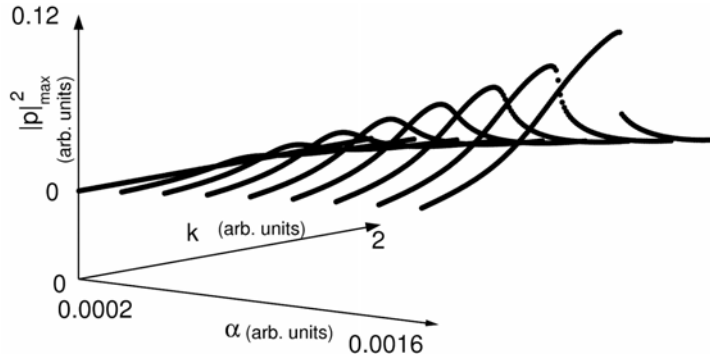


FIGURA 1. Amplitud del sonido vs.  $k$  para diferentes valores del coeficiente de acople  $\alpha$ . A medida que aumenta el acople, la amplitud en la resonancia crece. Para  $\alpha$  suficientemente grande, se establece una región de coexistencia, donde dos soluciones oscilatorias con amplitudes diferentes son posibles. Los parámetros utilizados para la integración numérica son  $(\gamma, p_s, \beta, c, r) = (7000\pi, 0,1, 0,01, 1 \times 10^{-6}, 0,71)$  en unidades adimensionales,  $\tau = 1,43 \times 10^{-4}s$ .

Cada punto representa el valor de  $|p|_{max}^2$ , calculado para una grilla de diferentes condiciones iniciales  $(x_0, y_0)$ . Cuando el acoplamiento es considerado, los valores de  $k$  para los que la amplitud del sonido es máxima se desplazan, y sus correspondientes valores pico aumentan. Tanto el corrimiento como la enfatización de los picos se observan incluso para valores de  $\alpha$  más pequeños que los necesarios para que ocurra la bifurcación que conduce a la coexistencia. En consecuencia, los fenómenos asociados al acoplamiento fuente-filtro tienen más probabilidades de ser identificados en la amplitud de las vocalizaciones, ya que su traza en las propiedades acústicas de las soluciones no se restringe a una región limitada en el espacio de parámetros. Cuando el acople es lo suficientemente fuerte, aparece una región de coexistencia, como es evidente en la figura 2. En esta figura, las amplitudes al cuadrado para

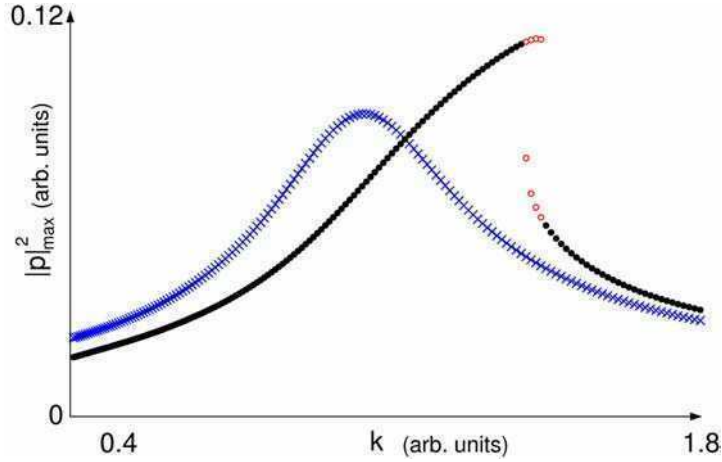


FIGURA 2. Detalle de la amplitud del sonido contra  $k$  y comparación con el sistema fuente-filtro desacoplados. Cruces representan amplitud de sonidos originados en un sistema fuente-filtro. Los puntos representan los sonidos generados con el tracto acoplado a la fuente. Puntos vacíos resaltan la región de coexistencia de soluciones periódicas: para los valores de  $k$  en esa región, la amplitud del sonido podrá ser uno de dos valores, dependiendo de las condiciones iniciales. Los parámetros utilizados para la integración numérica son  $(\gamma, p_s, \beta, c, r, \alpha) = (7000\pi, 0,1, 0,01, 1 \times 10^{-6}, 0,71, 0,0014)$  en unidades adimensionales,  $\tau = 1,43 \times 10^{-4} s$ .

un  $\alpha$  particular se grafican para el sistema acoplado, junto con los resultados del cálculo de la misma magnitud cuando se desprecia la contribución de la presión supra-siríngea en el término que fuerza la fuente. Cuando la hipótesis de separación fuente-filtro se supone válida, no hay bifurcación que conduzca a coexistencia de soluciones. Por otra parte, ningún cambio se observa en la frecuencia a la que ocurre el pico, y su amplitud es menor que su correspondiente cuando se considera el acople.

## 2. La aproximación

Las observaciones numéricas motivan trabajo analítico. Con el fin de averiguar el origen dinámico de los efectos introducidos por el acople fuente-tracto en la amplitud del sonido, vamos a reducir el sistema (4, 5) a un conjunto más simple de ecuaciones que preserven la dinámica de la amplitud de las soluciones oscilatorias.

Aquí proponemos estudiar el límite débilmente no lineal del sistema, en el que las no-linealidades introducen una desviación respecto de las soluciones periódicas de la parte lineal del sistema [Str00, EG08]. En este límite se pueden obtener expresiones analíticas para las desviaciones que se mencionan, tanto en la amplitud como en la fase de las nuevas oscilaciones. Diferentes soluciones oscilatorias aparecerán cuando se introduzcan las perturbaciones no lineales. Asimismo, los tipos de soluciones que se encuentren

cuando las no linealidades incluyan el acople con el tracto pueden ser diferentes a los encontrados cuando las no-linealidades contienen sólo las propias a la dinámica de la fuente.

Por estas razones, nos concentramos en el sistema

$$\ddot{x} + \omega_0^2 x = \mu \left[ \dot{x} - \tilde{c}x^2 \dot{x} + (1 - \dot{x})\tilde{\alpha} \sum_{n=0}^{\infty} (-r)^n x(t - n\tau) \right], \quad (21)$$

que es esencialmente el mismo que el descrito por las ecuaciones (4, 5), con  $\gamma = 1$ ,  $\mu = p_s - \beta$ ,  $\tilde{c} = c/\mu$ ,  $\tilde{\alpha} = \sqrt{p_s}\alpha/\mu$ ,  $\omega_0^2 = k$ .

**2.1. El límite débilmente no lineal.** La ventaja de estudiar el sistema (21) en el límite débilmente no lineal ( $\mu \ll 1$ ) es que, como la parte no lineal es tomada como una perturbación, las trayectorias serán desviaciones de oscilaciones armónicas. Se pueden encontrar condiciones para que esas trayectorias sean periódicas. Por otra parte, se pueden encontrar bifurcaciones que lleven, por ejemplo, a coexistencia de soluciones oscilatorias. La ecuación (21) puede escribirse en forma equivalente como

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -\omega_0^2 x + \mu [y - \tilde{c}x^2 y + (1 - y)\tilde{\alpha} \sum_{n=0}^{\infty} (-r)^n x(t - n\tau)]. \end{cases} \quad (22)$$

Antes de realizar aproximaciones, comenzamos por proponer el cambio de variables

$$\begin{cases} x = ae^{i\omega t} + \bar{a}e^{-i\omega t} \\ y = i\omega ae^{i\omega t} - i\omega \bar{a}e^{-i\omega t}, \end{cases}$$

donde  $a = a(t)$  es una nueva variable compleja y  $\bar{a}$  su compleja conjugada. Cambiando variables en el sistema (22) derivamos una ecuación para  $\dot{a}$ :

$$2i\omega \dot{a} = (\omega^2 - \omega_0^2) (a + \bar{a}e^{-2i\omega t}) + \mu f(a, \bar{a}, e^{i\omega t}, e^{-i\omega t}). \quad (23)$$

Hasta aquí no se han hecho aproximaciones, y esta última ecuación es idéntica a (22).

Si  $\mu = 0$ , el sistema tiene un punto fijo en  $\omega = \omega_0$ , donde  $a$  es constante y las oscilaciones son armónicas. Cuando consideramos el caso en el que la disipación y las no-linealidades son pequeñas ( $\mu \ll 1$ ), las soluciones de  $\omega \approx \omega_0$  se desvían ligeramente de oscilaciones armónicas y  $\dot{a} \ll a\omega$ . De esta manera, el cambio en  $a$  durante un período de oscilación  $T$  puede ser despreciado. Bajo estas suposiciones, muchos términos en el sistema pueden ser eliminados, dando lugar a una ecuación más simple que preserva la dinámica. La ecuación (23) se puede escribir en la forma

$$\dot{a} = \frac{\mu}{2i\omega} \sum_{n=0}^{\infty} F_n(a, \bar{a}) e^{i\omega n t},$$

donde el lado derecho de la ecuación es la serie de Fourier del lado derecho de (23), con coeficientes

$$F_n = \frac{1}{T} \int_t^{t+T} \left[ \frac{\omega^2 - \omega_0^2}{\mu} (a + \bar{a}e^{-2i\omega t'}) + \mu f(a, \bar{a}, e^{i\omega t'}, e^{-i\omega t'}) \right] e^{-in\omega t'} dt'.$$

Esta expansión es exacta si  $a$  es constante. Si  $a$  cambia lentamente,  $a = a(\mu t)$ , su cambio durante un periodo de oscilación  $T$  es pequeño, y podemos considerarla aproximadamente constante en la evaluación de la integral.

Puesto que sólo estamos interesados en cambios lentos en  $a$ , mantenemos sólo los términos no oscilantes en la expansión. Por medio de este procedimiento estándar, se eliminan todos los términos no resonantes en la ecuación, que son aquellos que tienen promedio nulo en un ciclo de oscilación de las variables lentas  $a$ . Después de esto y con la introducción de las nuevas variables  $t \rightarrow t' = (\mu/2)t$  y  $a \rightarrow A = a\sqrt{\tilde{C}}$ , se obtiene una ecuación sin dimensiones que describe la dinámica del desvío de las trayectorias respecto de oscilaciones armónicas,  $A(t') = \rho(t')e^{i\phi(t')}$ :

$$\begin{cases} \dot{\rho} &= \rho \left[ 1 + \frac{\alpha\sqrt{p_s}r}{(p_s - \beta)\omega} \frac{\sin(\omega\tau)}{1 + r^2 + 2r \cos(\omega\tau)} \right] - \rho^3 \\ \dot{\phi} &= \frac{\omega_0^2 - \omega^2}{(p_s - \beta)\omega} + \frac{\alpha\sqrt{p_s}}{(p_s - \beta)\omega} \frac{1 + r \cos(\omega\tau)}{1 + r^2 + 2r \cos(\omega\tau)}. \end{cases} \quad (24)$$

La búsqueda de soluciones para el sistema oscilatorio (4, 5) ahora se reduce a la búsqueda de puntos fijos en (24). Por lo tanto, observando las bifurcaciones en los puntos fijos de  $(\rho, \phi)$  podemos encontrar cambios cualitativos en el comportamiento oscilatorio de  $x$ , y de este modo del cuadrado de la amplitud por unidad de  $\alpha$  del sonido producido por el sistema fuente-tracto,

$$|p|_{max}^2 = (1 - r)^2 \frac{p_s \rho^{*2}}{1 + 2r \cos(\omega\tau) + r^2}, \quad (25)$$

donde  $\rho^*$  es el valor de  $\rho$  en el punto fijo:

$$\rho^{*2} = 1 + \frac{\alpha\sqrt{p_s}r}{(p_s - \beta)\omega} \frac{\sin(\omega\tau)}{1 + r^2 + 2r \cos(\omega\tau)},$$

con  $\omega$  satisfaciendo

$$\omega_0^2 = g(\omega) = \omega^2 + \alpha\sqrt{p_s} \frac{1 + r \cos(\omega\tau)}{1 + r^2 + 2r \cos(\omega\tau)}. \quad (26)$$

Con estas expresiones podemos identificar las condiciones bajo las que existen tres, uno, o ningún punto fijo para  $\omega \approx \omega_0$ .

Por otra parte, la estabilidad de la solución se puede determinar analíticamente. Estos resultados se resumen en la figura 3. Dado un valor de  $\omega$  que satisfice la condición (26), el sistema presentará un punto fijo estable si  $\frac{\partial g}{\partial \omega} > 0$ , inestable si  $\frac{\partial g}{\partial \omega} < 0$ . La función  $g(\omega)$  es creciente desde  $\omega = 0$ , de manera que los puntos fijos serán estables a menos que suceda que  $g(\omega)$  presente un máximo en un valor que satisfaga  $\frac{\partial g}{\partial \omega} = 0$ ,  $\frac{\partial^2 g}{\partial \omega^2} < 0$ . En este máximo ocurre una bifurcación nodo-silla y coexisten tres puntos fijos: dos estables y uno inestable. Si se encuentra tal máximo,  $g(\omega)$  también presentará un mínimo para un valor mayor de  $\omega$  ( $\frac{\partial g}{\partial \omega} = 0$ ,  $\frac{\partial^2 g}{\partial \omega^2} > 0$  en ese punto). Un punto fijo estable y uno inestable se encontrarán en una nueva bifurcación nodo-silla. La existencia de este par de bifurcaciones nodo-silla requiere que  $g(\omega)$  presente un máximo. Por lo tanto, es posible encontrar, para  $(r, \omega)$  fijos, el menor valor de  $\alpha$  para el que las condiciones (26) y  $\frac{\partial^2 g}{\partial \omega^2} > 0$  se pueden satisfacer. Estos valores críticos de  $(\alpha, \omega_0)$  se encuentran donde ocurre una bifurcación de tipo *cusp*. Para cada punto fijo de  $a$ , el sistema fuente-filtro oscila.

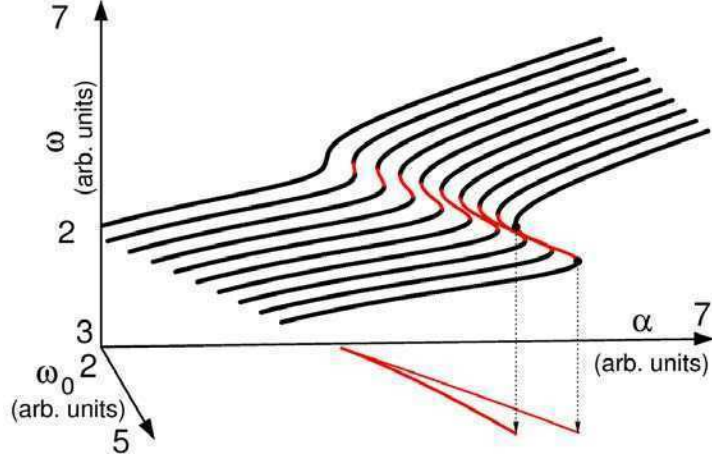


FIGURA 3. Frecuencia angular para puntos fijos del sistema (24). Las líneas sólidas oscuras indican puntos fijos estables, donde se cumple la condición (26) y  $\frac{\partial g}{\partial \omega} > 0$ . Líneas más delgadas y claras indican puntos fijos inestables, en los que  $\frac{\partial g}{\partial \omega} < 0$ . Las líneas en el plano  $(\alpha, \omega_0)$  delimitan la región de coexistencia de puntos fijos. Sobre esas líneas  $\frac{\partial g}{\partial \omega} = 0$ , indicando la ocurrencia de bifurcaciones nodo-silla de puntos fijos. El punto en el que se encuentran es donde ocurre una bifurcación de tipo *cusp*: ese es el valor crítico para  $(\alpha, \omega_0)$  en el que es posible la coexistencia. Los parámetros utilizados para la integración numérica son  $(p_s, \beta, r, \tau) = (5, 1, 0, 1, 0, 51, 1, 0)$  en unidades adimensionales.

### 3. Efectos dinámicos observables en la amplitud

En la figura 4 graficamos los cuadrados de las amplitudes de las oscilaciones por unidad de  $\alpha$  (calculados a partir de la ecuación (25)) correspondientes a los puntos fijos que se muestran en la figura 3. En la búsqueda de trazas de acople fuente-tracto en la amplitud del sonido generado por el sistema completo, se comparan los resultados anteriores con los obtenidos al suponer independencia del tracto de la dinámica de la fuente. En este último caso, los puntos fijos de  $a$  ocurren en  $(\rho^* = 1, \omega = \omega_0)$ . Consecuentemente, no existen bifurcaciones que conduzcan a una coexistencia de ciclos límite.

Además de esto, la amplitud por unidad de  $\alpha$  de las oscilaciones no varía con  $\alpha$ . Las amplitudes resultantes se muestran en el panel superior de la figura 5. En el panel central de la figura se muestran los cuadrados de las amplitudes, calculados para los mismos valores de los parámetros que para el sistema con acople acústico. Junto a ellos, se muestran las líneas de bifurcación de tipo *cusp* de la misma manera que discutimos anteriormente.

También se incluye en este panel una línea de puntos para ilustrar otro de los fenómenos originados por el acoplamiento: el corrimiento del umbral de fonación. La existencia de ciclos límite en el sistema acoplado requiere

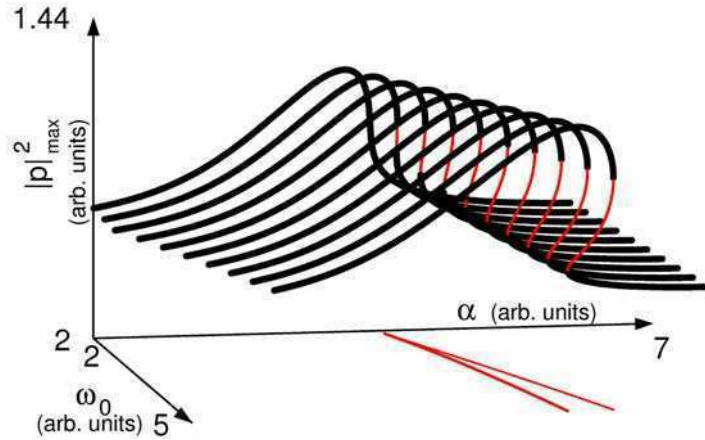


FIGURA 4. Amplitud de sonido para soluciones estacionarias. Las líneas sólidas oscuras indican puntos fijos estables, donde se cumple la condición (26) y  $\frac{\partial g}{\partial \omega} > 0$ . Líneas más delgadas y claras indican puntos fijos inestables, en los que  $\frac{\partial g}{\partial \omega} < 0$ . Las líneas en el plano  $(\alpha, \omega_0)$  delimitan la región de coexistencia de puntos fijos. Sobre esas líneas  $\frac{\partial g}{\partial \omega} = 0$ , indicando la ocurrencia de bifurcaciones nodo-silla de puntos fijos. El punto en el que se encuentran es donde ocurre una bifurcación de tipo *cusp*. Los valores de los parámetros son los mismos que en la figura 3.

que se satisfaga la ecuación (26). Para un conjunto fijo de  $(\alpha, p_s, r, \tau)$ , los valores de  $\omega_0$  debajo de cierto umbral no dan lugar a oscilaciones. Estos valores umbral se calculan y se representan como una línea de puntos en el plano  $(\alpha, \omega_0)$ . El panel inferior de la figura ilustra las diferencias cualitativas entre las amplitudes de sonidos que salen de un sistema de fuente-tracto acoplados y los que son producidos en un generador de sonido donde se asume válida la separación fuente-filtro. Se elige un valor de  $\alpha$  y se grafican las amplitudes al cuadrado por unidad de  $\alpha$  para ambas aproximaciones. El sistema acoplado muestra coexistencia de soluciones estables con diferente amplitud, corrimiento y enfatización de los picos de resonancia, y la introducción de un umbral de fonación dependiente de  $\alpha$ .

#### 4. Propiedades acústicas de las soluciones

El acople fuente-tracto introduce complejidad en el sistema de ecuaciones que describen la dinámica del órgano vocal. Aun teniendo en cuenta una dinámica simple para la fuente, el sistema acoplado fuente-tracto presenta un diagrama de bifurcaciones no trivial, que incluye una bifurcación de tipo *cusp* que conduce a la aparición de una región de coexistencia de soluciones periódicas. En la sección anterior, también encontramos mecanismos dinámicos que conducen a fenómenos que afectan a la amplitud de las soluciones oscilatorias, así como a un corrimiento en los valores de los parámetros necesarios para que ocurra la bifurcación de Hopf que da origen a las oscilaciones.

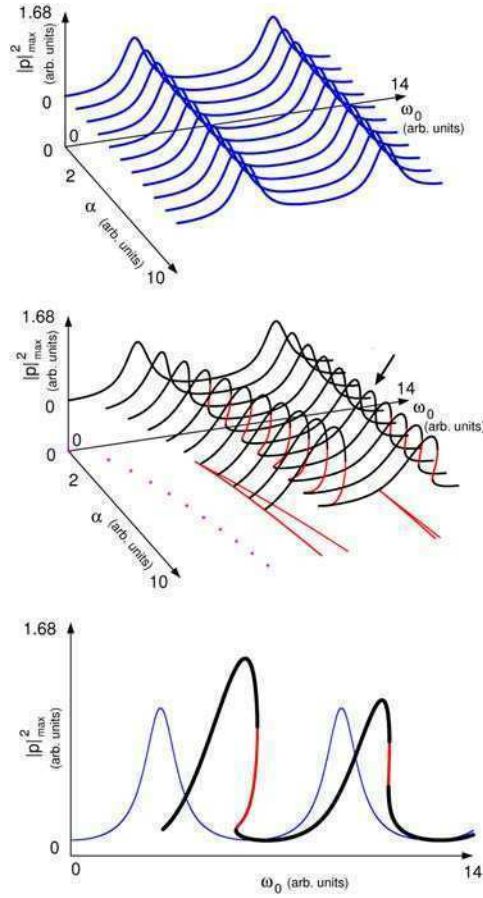


FIGURA 5. Amplitudes (al cuadrado, por unidad de  $\alpha$ ) de sonidos originados en el sistema fuente-filtro desacoplados comparadas con las de sonidos originados en el sistema fuente-tracto interactuantes. Sonidos provenientes de un sistema fuente-filtro desacoplados (panel superior). Los sonidos calculados a partir del sistema fuente-tracto interactuantes presentan regiones de coexistencia de soluciones estables. Las líneas gruesas y oscuras representan soluciones estables; las líneas más claras y finas indican soluciones inestables. Las regiones de coexistencia de soluciones estables se indican con las líneas en el plano  $(\alpha, \omega_0)$  que se intersectan entre sí. Los umbrales de fonación son indicados por la línea de puntos en el plano  $(\alpha, \omega_0)$  (panel central). Comparación de amplitudes de sonidos producidos por los dos sistemas con  $\alpha = 7,0$ , según lo indicado por la flecha en el panel del medio (panel inferior). Los valores de los parámetros son los mismos que en la figura 3.

Las características de las soluciones oscilatorias dependen de parámetros que dan cuenta de variables fisiológicas del sistema. De esta manera, la complejidad de las soluciones, que deja su huella en las propiedades acústicas

---

de las vocalizaciones producidas, puede ser controlada en el sistema fuente-tracto por gestos motores simples adecuadamente coordinados.

La actividad del músculo *siringealis ventralis* ( $vS$ ) está directamente relacionada a la frecuencia fundamental del sonido producido en la siringe [GS95], en un modo que conduce a la hipótesis de que esta variable fisiológica es responsable del control activo de la rigidez de los labios (parámetro  $k$  en la ecuación (4)) [LGM02, SAGM10]. Además, se han reportado casos en los que la actividad del tracto vocal está coordinada con la frecuencia de las vocalizaciones [Now87, RSFB06]. Más allá de la hipótesis de separación fuente-filtro, vocalizaciones con propiedades acústicas no triviales pueden ser generadas por caminos simples en el espacio de parámetros que dan cuenta de los gestos motores que controlan la tensión labial  $k$ , la presión sub-siringea  $p_s$  y la longitud del tracto vocal  $L$ .

La coordinación de las actividades del tracto vocal y la siringe es uno de los mecanismos por los cuales las aves oscinas enfatizan la frecuencia fundamental de la vocalización. Mediante el ajuste de la longitud  $L$  de la tráquea, la frecuencia del sonido producido por las vibraciones de los labios en la siringe se hace corresponder con la resonancia del tubo con el que se filtra el sonido  $\nu = v_s/(4L)$ . La introducción del acople fuente-filtro en el modelo predice, para ciertos valores de los parámetros, una enfatización de este efecto: los sonidos resonantes en el sistema acoplado presentan una mayor amplitud que en la aproximación fuente-filtro. Para construir la figura 2, se generaron sonidos sintéticos con sistemas idénticos, uno de ellos acoplado, y el otro desacoplado. Para un  $L$  fijo, la frecuencia de la oscilación generada en los labios fue variada (mediante un barrido de los parámetros que dan cuenta de la rigidez de los labios  $k$ ), y se computó la amplitud de las oscilaciones. En las simulaciones del sistema fuente-tracto con acople, las resonancias ocurren a frecuencia fundamental mayor y son más pronunciadas.

Para ilustrar las diferencias en el contenido espectral de las vocalizaciones producidas en ambas aproximaciones, sintetizamos vocalizaciones mediante integración numérica del sistema (4, 5) (se integró una versión modificada de la ecuación (4) con  $p_i = 0$  para producir las vocalizaciones del sistema desacoplado fuente-filtro). Se introduce un gesto de presión muy simple, que consiste simplemente de un aumento por encima del valor en el que los tejidos comienzan a oscilar y, después de un intervalo de tiempo, un retorno a nivel sub-umbral. Durante ese tiempo tiene lugar la vocalización. La longitud del tracto se mantiene fija, ajustada de forma que su frecuencia de resonancia coincida con la frecuencia fundamental del sonido producido por la fuente sola. Las vocalizaciones son representados en la figura 6 (presión acústica y sonograma).

Es pertinente destacar que la energía del sistema fuente-tracto acoplados se concentra en la frecuencia fundamental  $f_0$  y en todos los armónicos ( $f_0, 2f_0, 3f_0, \dots$ ). En contraste, los picos de energía en la sílaba fuente-filtro se producen como se espera en un sonido filtrado por un tubo abierto-cerrado con  $L = v_s/(4f_0)$ , es decir, en ( $f_0, 3f_0, \dots$ ). El término cuadrático introducido por la realimentación en el forzante del sistema (4) es responsable de este efecto.



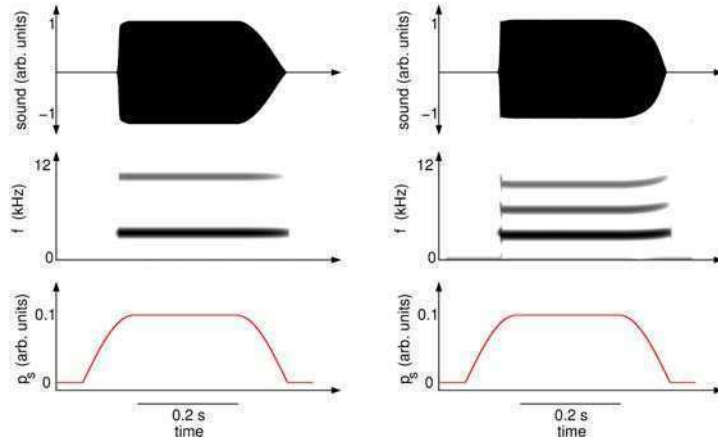


FIGURA 6. Vocalizaciones sintéticas producidas mediante integración numérica de las ecuaciones (4, 5), considerando independencia fuente-filtro (paneles de la izquierda) y la fuente-tracto acoplados (paneles de la derecha). Se produce sonido (paneles superiores) cuando el sistema es conducido por un gesto de presión simple (paneles inferiores). Los sonogramas (paneles del medio) muestran que el contenido espectral de las vocalizaciones provenientes de uno y otro sistema son cualitativamente diferentes. Los parámetros utilizados para la integración numérica son  $(\gamma, \beta, c, r, \alpha) = (7000\pi, 0,01, 1 \times 10^{-6}, 0,71, 0,0014)$  en unidades adimensionales,  $\tau = 1,43 \times 10^{-4}s$ .

Se han realizado experimentos en algunas especies para estudiar de qué forma el tracto vocal modifica el sonido generado en la fuente, en los que el ambiente donde las aves cantan está lleno de *heliox* (una mezcla de helio y oxígeno) [Now87, BPDH97]. El cambio de ambiente produce un cambio en la velocidad del sonido, permitiendo llevarla hasta  $550 m/s$ . Si la fuente y el tracto no interactúan, las frecuencias fundamentales de la vocalizaciones grabadas no deberían modificarse con el aumento en velocidad del sonido. Mediante la simulación de un experimento de heliox con sonidos sintéticos generados por un sistema acoplado fuente-tracto, esperamos ser capaces de hacer predicciones cuantitativas sobre los cambios en la frecuencia fundamental y el contenido espectral a observar cuando velocidad del sonido aumenta.

Con esto en mente, generamos una serie de sonidos sintéticos con todos los parámetros del sistema (4, 5) fijos, excepto por  $v_s$ , que se varió en un rango de  $341 m/s$  a  $520 m/s$ . La frecuencia fundamental observada en los sonidos originados en el sistema no acoplado se mantuvo constante durante el experimento simulado, tal como se esperaba. No se observó cambio en la frecuencia fundamental o en las resonancias superiores, sino un cambio en la relación de los picos (ver figura 7, paneles de la izquierda). Los sonidos sintetizados utilizando el sistema acoplado fuente-tracto aumentan su frecuencia fundamental a medida que la velocidad del sonido se incrementa. Asimismo, el corrimiento en los picos de amplitud tiende a ser más grande cuanto

mayor es la frecuencia a la que se presentan (ver figura 7, paneles de la derecha). Se han observado pequeños corrimientos en las frecuencias cuando

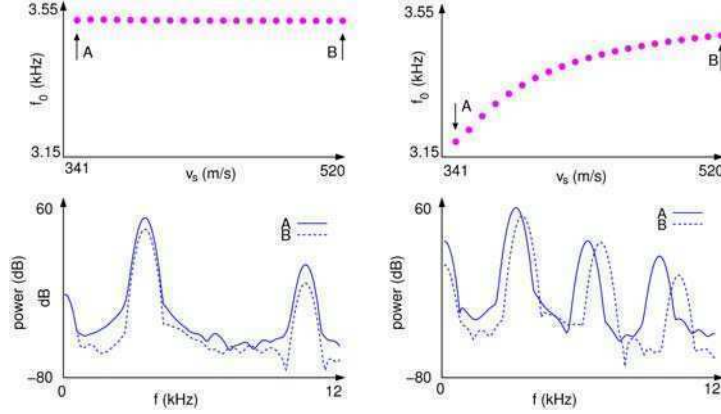


FIGURA 7. Frecuencias fundamentales de vocalizaciones sintéticas y sus espectros de energía para un medio de densidad variable. Las frecuencias fundamentales de los sonidos no cambian con la velocidad del aire si la fuente y el filtro no interactúan (panel superior izquierdo). Cuando la fuente está acoplada con el tracto, hay una deriva en la frecuencia fundamental (panel superior derecho). Las flechas *A* y *B* en los paneles superiores indican dos densidades del aire diferentes, para las que se calculó el espectro de energía de los sonidos sintéticos (paneles inferiores). cuando la densidad del aire cambia, el espectro del sonido sintético del sistema fuente-filtro sólo presenta cambios en los valores relativos entre los picos de resonancia (panel inferior izquierdo); mientras que en los picos de resonancia de las vocalizaciones del sistema fuente-tracto acoplados se aprecia un corrimiento, que se incrementa al ir a armónicos superiores (panel inferior derecho). Los parámetros utilizados para la integración numérica son  $(\gamma, \beta, c, r, \alpha) = (7000\pi, 0,01, 1 \times 10^{-6}, 0,71, 0,0014)$  en unidades adimensionales,  $\tau = 9,6cm/v_s$ .

la densidad del aire se reduce al mezclarlo con heliox [Now87]. Aquí hemos demostrado que, incluso con una descripción simple de la dinámica de las membranas oscilantes en la siringe, es de esperar que ocurra un cambio en la frecuencia ante un cambio en la densidad de la atmósfera, y que de este efecto es responsable la realimentación retardada introducida por el acople de la fuente con el tracto.

## 5. Discusión

En esta parte hemos estudiado la dinámica de un sistema fuente de sonido-tracto interactuantes. Nuestro modelo consiste de un oscilador acoplado a un tubo simple. Los tractos vocales son en realidad mucho más complejos e incluyen, por ejemplo, el pico y la cavidad oro-faríngea-esofágica (*OEC*), cuyas actividades han visto coordinadas con la de la fuente de

---

sonido [RSFB06, FRS06]. En general, sin embargo, ellas introducen sólo un filtrado al sonido, y sus contribuciones a la realimentación son despreciables. Encontramos que cuando las frecuencias de los sonidos generados en la fuente están cerca de las frecuencias de resonancia del tubo, ocurre un corrimiento y una enfatización en las resonancias del sistema. Para algunos valores de los parámetros, también encontramos coexistencia de soluciones periódicas estables. Esto implica la posibilidad de que ocurran cambios repentinos en la salida acústica del sistema, tanto en amplitud y frecuencia, incluso ante cambios suaves en los parámetros.

Las vocalizaciones complejas producidas por los oscinos ocurren como resultado de la interacción de un sistema nervioso y una periferia biomecánica. En esta parte, contribuimos a la tarea de determinar dónde se origina tal complejidad. Este tema ha sido abordado en [ZRS08, ML05, ZM04]. La hipótesis es que la complejidad de las vocalizaciones no requiere control neural activo: que bien se puede lograr mediante la conducción de una periferia altamente no lineal, con instrucciones motoras simples. Así como en la siringe ocurre física no lineal que introduce riqueza espectral en los sonidos, las interacciones acústicas entre la fuente de sonido y el tracto agregan complejidad. El mecanismo propuesto aquí apoya la idea de que, con el fin de comprender la complejidad del canto de los pájaros, es necesario estudiar en paralelo el sistema central de control neuronal y la dinámica de la periferia.

## Parte 3

# Interfaces cerebro-máquina basados en los modelos

El comportamiento motor complejo emerge de las interacciones entre el sistema nervioso central y los sistemas efectores periféricos [CB97]. Esta interacción se ilustra claramente en la producción de canto de las aves, donde un dispositivo altamente no lineal es capaz de generar una variedad de sonidos diferentes, incluso cuando es conducido por instrucciones fisiológicas relativamente simples [SAGM08, AM08]. Recientemente, el modelado del órgano vocal aviar ha ayudado a entender las relaciones entre distintas propiedades acústicas del canto, que no están bajo control neuronal directo, sino que están determinadas por la biomecánica del sistema periférico [SAGM08, AM08].

En esta parte desarrollamos el avance hacia la construcción de un dispositivo electrónico que lee continuamente las instrucciones fisiológicas que conducen la siringe, e integra las ecuaciones del modelo que domina su dinámica en el lapso de tiempo entre lecturas. En un programa que esperamos sea paradigmático, basamos el dispositivo en el modelado matemático de baja dimensión de los procesos físicos que ocurren en la periferia para originar las vocalizaciones.

Esta forma presenta una ventaja respecto de otros acercamientos al problema de las interfaces cerebro-máquina (BMI y BCI por sus siglas en inglés: *Brain Machine Interfaces* y *Brain Computer Interfaces*), en los que registros neurofisiológicos de alta dimensión son procesados estadísticamente para encontrar características que codifiquen determinadas intenciones motoras. Por un lado, la comprensión del comportamiento como el resultado de la interacción entre un sistema nervioso central y un sistema periférico complejo, implica que la codificación de las instrucciones motoras no se busca en un conjunto de relaciones causales entre los registros fisiológicos y el comportamiento. En cambio, se trata de encontrar los parámetros que determinan distintos comportamientos en un sistema complejo, que puedan ser relacionados con instrucciones fisiológicas de control. Por otro lado, al alcanzar un nivel adecuado de descripción del aparato fonador en términos de un sistema de ecuaciones diferenciales de baja dimensión, la computación del modelo es suficientemente sencilla como para efectuarse en procesadores de señales digitales portátiles de uso estándar.

La especie más estudiada de ave oscina es el diamante mandarín o *Zebra finch* (*Taeniopygia guttata*), cuyo canto se compone de 3-8 sílabas distintas con una variedad de características acústicas. Mientras que muchas especies de aves oscinas producen sonidos con bajo contenido de armónicos superiores (tonales), el canto del diamante mandarín se compone de sílabas espectralmente ricas tanto como tonales. En los próximos capítulos presentamos el desarrollo en modelado matemático del aparato fonador del Diamante mandarín, el análisis del modelo en su complejidad, la relevancia de la complejidad para reproducir las características del canto y la pertinencia fisiológica de los parámetros de control. Este desarrollo converge a la construcción de un dispositivo que, conducido por un pájaro a través de sus propios gestos motores, genera canto sintético realista en tiempo real.

## Comportamiento motor complejo: el canto del diamante Mandarín

El primer paso para la construcción de un dispositivo que emule comportamiento es comprenderlo. Orientamos la comprensión hacia el modelado de la biofísica de la periferia biomecánica que lo genera y de los mecanismos de interacción con el generador central de patrones motores.

En este capítulo se introduce el canto del diamante mandarín, como modelo animal de un comportamiento motor aprendido complejo. También se introduce un modelo de aleteo extendido que ha tenido éxito al capturar los orígenes dinámicos tras la producción de sonidos de propiedades acústicas diversas como la del diamante mandarín. Desde que fue presentado por miembros de Laboratorio de Sistemas Dinámicos [SAGM08], ha servido para explorar los mecanismos de control periférico involucrados en la producción de canto del diamante mandarín. En paralelo, y por su capacidad de reproducir canto sintético realista cuando es conducido por caminos adecuados en un espacio de parámetros fisiológicamente relevante y de baja dimensión, es la base de un dispositivo bioprotésico [SAGM10, APGM12].

### 1. Canto del diamante Mandarín

El diamante mandarín tiene un canto estereotipado, tanto en ritmo (cómo construye la secuencia de sílabas) como en propiedades acústicas (cómo son los sonidos que emite) [ZM04].

Una *frase* de canto contiene un número (variable) de notas introductorias seguido por la repetición de un *motivo*. Las notas introductorias son sílabas que duran alrededor de 100 *ms.* y tienen baja frecuencia (400 – 600 *Hz*). El motivo es una secuencia de 2 – 8 sílabas, formadas por sonidos que cubren un rango amplio de dos características acústicas: la frecuencia fundamental y el contenido espectral.

Entre las sílabas que forman el motivo, las hay de alta frecuencia fundamental y bajo contenido espectral, y de bajo contenido espectral y alta frecuencia fundamental. El contenido espectral determina una característica perceptible del sonido tan importante como el timbre. En [SAGM08] se introdujo un índice de contenido espectral ( $I_{SC}$ ) para explorar cuantitativamente el rango cubierto por la diversidad de sonidos del diamante mandarín en el espacio de propiedades acústicas ( $f_0, I_{SC}$ ) (frecuencia fundamental, índice de contenido espectral).

El índice le asigna un valor numérico a la dispersión de la energía de un segmento de sonido a lo largo de sus armónicos. Es un estimador del primer momento la distribución de energía en las componentes espectrales, relativo

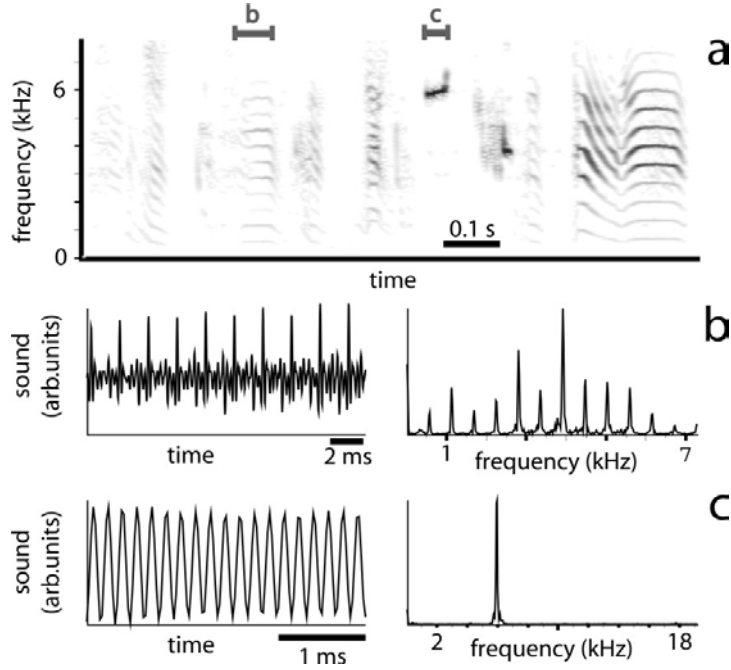


FIGURA 1. Una frase de canto de diamante mandarín, con una visualización de la serie temporal y FFT de un par de segmentos ejemplares. En el sonograma del canto se aprecia la diversidad en frecuencia y contenido espectral de las vocalizaciones (a). El análisis espectral del segmento indicado con la letra b, muestra que su frecuencia fundamental es baja ( $400 - 600 \text{ Hz}$ ) espectralmente rico, porque su energía está distribuida a lo largo de sus armónicos (b). El segmento marcado por la letra c es de alta frecuencia (algunos  $\text{kHz}$ ) y tonal, o espectralmente pobre, con la energía concentrada en la frecuencia fundamental (c). (Extraído de [SAGM08]).

a la frecuencia fundamental:

$$I_{SC} = \frac{1}{f_0} \frac{\sum_i f_i e_i}{\sum_i e_i}. \quad (27)$$

Para calcularlo, se aplica la transformada rápida de Fourier (FFT) al segmento de sonido, y se calcula la energía  $e_i$  en cada frecuencia  $f_i$  como el cuadrado del módulo de esa componente.

La exploración del rango de propiedades acústicas cubierto por el diamante mandarín se llevó a cabo tomando segmentos de canto provenientes de diversas sílabas de distintos sujetos, encontrando su frecuencia fundamental y calculando su índice de contenido espectral. Lo que se encontró es que, del espacio de propiedades acústicas de dos dimensiones ( $f_0, I_{SC}$ ), los sonidos que emite el diamante mandarín ocupan apenas una banda, concentrada estrechamente alrededor de una línea (que puede ser ajustada por una función exponencial decreciente [SAGM08]). La relación entre frecuencia fundamental y contenido espectral de los sonidos emitidos por el diamante mandarín

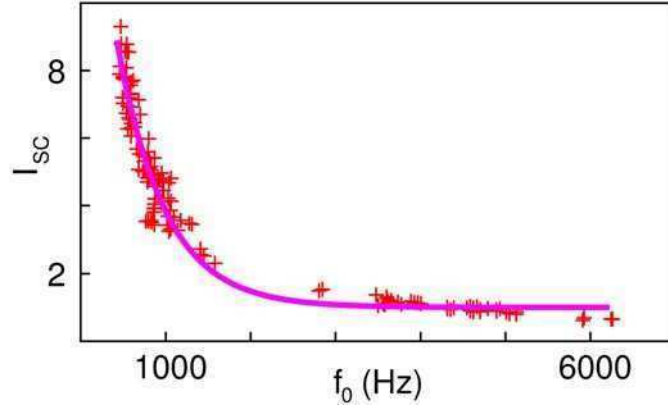


FIGURA 2. Hay una relación funcional entre la frecuencia fundamental y el índice de contenido espectral para los cantos del diamante mandarín. En el gráfico, los puntos representan segmentos de 512 muestras de canto (a una frecuencia de muestreo de  $44,1\text{ KHz}$  de cantos de 5 animales distintos. Para cada segmento, se encontró la frecuencia fundamental y se calculó el índice de contenido espectral. La línea sólida es un ajuste del tipo  $f(x) = Ae^{-Cx} + 1$ .

queda bien determinada por una relación funcional: determinada una, queda unívocamente determinada la otra.

## 2. Modelo del órgano fonador

La diversidad de características acústicas presentes en el canto del diamante mandarín puede ser explicada por la dinámica que exhibe el modelo matemático de su órgano vocal [LGM02, SAGM08]. Este modelo de baja dimensión, al ser conducido por el gesto apropiado en el espacio de parámetros, es capaz de producir canto sintético comparable con el real. Mediante la implementación de este modelo en un procesador de señales digitales, podremos de construir un órgano vocal bio-protésico.

**2.1. La siringe.** Describimos la dinámica de las membranas en la siringe del diamante mandarín mediante una versión extendida del modelo de aleteo. Este presenta un buen compromiso entre el nivel de descripción y la dimensión del sistema. En [SAGM08], se propuso una versión del modelo de aleteo extendido en el que pueden ocurrir dos bifurcaciones diferentes, originando en los labios oscilaciones de características diversas.

Del mismo modo que en las versiones anteriores del modelo extendido, se realiza una descripción cinemática del perfil de la válvula siríngica en términos del desplazamiento del equilibrio la posición central de cada labio  $x$ . El movimiento de la posición del punto medio de los labios obedece la segunda ley de Newton:

$$\begin{cases} \frac{dx}{dt} = y, \\ \frac{dy}{dt} = (1/m) [-k(x)x - b(y)y - cx^2y + a_1 p_{av}]. \end{cases} \quad (28)$$



---

En el lado derecho de la segunda ecuación, el primer término describe la restitución no lineal elástica de los labios ( $k(x) = k_1 + k_2x^2$ ), el segundo término representa la disipación no lineal ( $b(y) = b_1 + b_2y^2$ ), y el tercero una saturación no lineal que limita el movimiento labial. El sistema es forzado por el último término, que representa la fuerza ejercida por la presión interlabial.

En el término de forzado  $p_{av}$  es el promedio espacial de la presión de conducción siringea calculado en la sección 2 del capítulo 2,

$$p_{av} = p_s \left(1 - \frac{a_2}{a_1}\right) = p_s \frac{x_{01} - x_{02} + 2\tau_f y}{x_{01} + x + \tau_f y}. \quad (29)$$

Notar que en este término se está despreciando la presión a la entrada del tracto frente a la presión sub-siringea ( $p_i = 0$ ). Esto es equivalente a suponer que se cumple la *hipótesis de separación fuente-filtro*. Poniendo las ecuaciones (28) y (29) juntas, se completa un modelo de aleteo extendido que rige la dinámica de los labios en la siringe del diamante mandarín.

En este modelo, características acústicas de las soluciones están determinadas por parámetros significativos fisiológicamente. Suponiendo que el coeficiente lineal en el término de la restitución del sistema (28) es proporcional a la tensión de los músculos siringeos ventrales (vS), este modelo es capaz de producir canto sintético realista.

**2.2. El tracto.** Bajo la hipótesis de separación fuente-filtro, el tracto vocal superior actúa como un filtro pasivo, modificando el contenido espectral del sonido producido en la fuente. En esta parte, donde se está introduciendo un modelo que sea lo suficientemente sencillo como para ser integrado en tiempo real, resaltamos, entre los elementos que componen el tracto vocal superior, la tráquea. Se puede generar sonido sintético con este modelo si la presión del sonido en la entrada del tracto se calcula como

$$p_i(t) = x(t) - r p_i(t - 2L/v_s), \quad (30)$$

donde  $L$  es la longitud de un tubo hace las veces de tráquea,  $c$  es la velocidad del sonido y  $r$  el coeficiente de reflexión de la onda sonora en el extremo del tubo abierto a la atmósfera. El sonido de salida se computa como la presión a la salida del tracto vocal,  $p_o(t) = (1 - r)p_i(t - L/v_s)$  [AM09, MGGS03]. Esta expresión refleja que la presión de salida es la onda de sonido transmitida al final del tracto.

Este modelo para el aparato fonador queda completo, con una fuente de sonido en la que se producen sonidos diversos mediante distintos mecanismos dinámicos, y un tracto vocal que actúa como filtro, terminando de determinar las propiedades acústicas de las soluciones del modelo.

**2.3. La dinámica del modelo.** El modelo matemático del órgano vocal representada por (28, 29 y 30), en lugar de un intento de obtener un filtro estadístico o un conjunto de reglas de causalidad entre un comando motor codificado y un comportamiento, es una descripción física del efector periférico complejo. Mediante el estudio de la dinámica no lineal de este modelo, especialmente la de sus soluciones oscilatorias, se encuentran relaciones no triviales entre trayectorias en el espacio de parámetros y características acústicas de los sonidos sintetizados por él.

---

En este modelo, el tracto vocal no-interactuante actúa como un filtro pasivo: sonidos producidos en la siringe se ven alterados por él sólo en el peso relativo de sus componentes armónicas. Es por esto que el análisis dinámico en busca de soluciones oscilatorias cualitativamente diferentes se hace en las ecuaciones que rigen la dinámica de la siringe. Una descripción detallada del conjunto de soluciones y bifurcaciones del sistema (28, 29) se llevó a cabo en [SAGM08, AM08], y un diagrama que sintetiza los hallazgos fue construido para el espacio de parámetros  $(k_1, p_s)$ . Este conjunto de parámetros es el más plausible de controlar las propiedades acústicas de los sonidos, puesto que se relaciona con la restitución de los labios que oscilan y la intensidad de la fuerza ejercida por el flujo de aire. Estos parámetros también se identifican con las instrucciones fisiológicas utilizadas por las aves para el control motor [LGM02, ML05, GC04].

La existencia de componentes no lineales en las fuerzas de la restitución de los labios juegan un papel importante en la generación de las oscilaciones de baja frecuencia y alto contenido espectral [SAGM08, AM08]. Dentro del escenario dinámico rico que presenta el sistema, dos mecanismos distintos que dan lugar a soluciones oscilatorias se destacan: una bifurcación de Hopf y una bifurcación nodo-silla en un ciclo límite (SNIC, por su sigla en inglés *Saddle-Node on an Invariant Cycle*). Ellas están ilustradas en la figura 3 junto con una representación esquemática de las soluciones y sus propiedades espectrales. A través de una bifurcación de Hopf sub-crítica, un punto fijo pierde estabilidad se crea y un pequeño ciclo límite estable [Str00]. Oscilaciones nacen de este modo con amplitud nula y frecuencia finita; por ende los sonidos producidos de esta manera son tonales. Cuando un punto fijo de tipo silla de montar (*saddle*) y un punto fijo no-silla colisionan en un ciclo límite, se produce una bifurcación SNIC en la que nacen oscilaciones de de amplitud finita y período infinito [Str00]. Los sonidos producidos por este mecanismo en consecuencia tienen una frecuencia baja y su contenido espectral es rico.

### 3. Una restricción de la biomecánica

La presentación de este modelo sencillo del órgano fonador persigue, entre otros, el objetivo de sintetizar canto comparable con el real. Manteniendo fijos los parámetros que dan cuenta de la geometría de la válvula siríngica y el tracto, y haciendo un barrido en el espacio de parámetros  $(k_1, p_s)$ , se produjeron esas síntesis.

Situándose en la región adecuada del espacio de parámetros, se logró reproducir sonidos de diverso contenido espectral y frecuencia fundamental, dentro de los mismos rangos de propiedades acústicas observados en los cantos estudiados. ¿Pero cómo tienen que ser las instrucciones motoras (los parámetros  $(k_1, p_s)$ ) para que se reproduzca, además, la restricción a un subespacio determinado por una relación funcional entre las propiedades acústicas  $(f_0, I_{SC})$ ?

Computando frecuencia y contenido espectral para los sonidos sintetizados durante el barrido en el espacio de parámetros, se obtiene un conjunto de puntos concentrados alrededor de la misma curva exponencial que describe la relación funcional entre las propiedades acústicas de los sonidos grabados

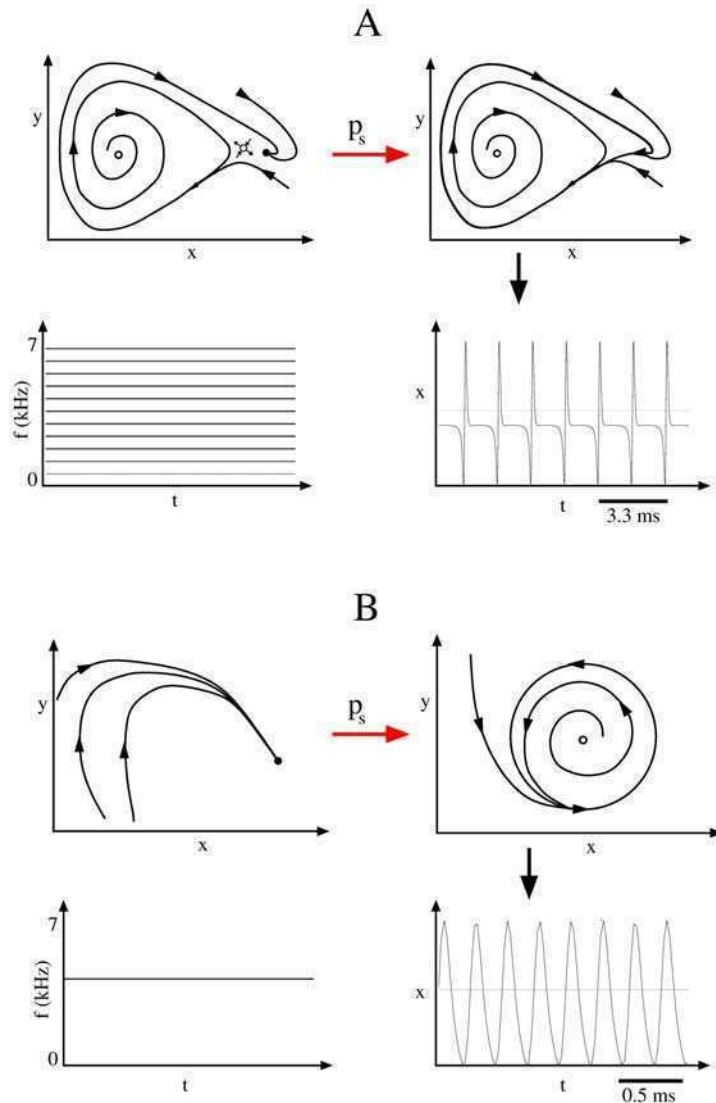


FIGURA 3. Esquema de las bifurcaciones que conducen a soluciones oscilatorias encontradas en el modelo de la siringe. Cuando el parámetro  $p_s$  atraviesa una bifurcación SNIC, dos puntos fijos colisionan en un ciclo límite (A). El ciclo límite nace con período grande y contenido espectral rico, como se ilustra en la evolución temporal de la variable  $x$  y su sonograma correspondiente. Cuando el parámetro  $p_s$  atraviesa una bifurcación de Hopf, un punto fijo estable pierde estabilidad frente a un ciclo límite naciente (B). Cerca de la bifurcación, el ciclo límite presenta período pequeño y contenido espectral tonal, como se ilustra en la evolución temporal de la variable  $x$  y su sonograma correspondiente. Ambas bifurcaciones ocurren en distintas regiones del espacio de parámetros  $(k_1, p_s)$  [SAGM08, AM08].

---

(figura 4). Es decir que la restricción de propiedades acústicas al subespacio determinado por una relación funcional entre contenido espectral y frecuencia fundamental viene impuesta por la dinámica de la fuente, aún cuando los gestos fisiológicos motores no estén limitados a una región determinada del espacio de parámetros que da cuenta de ellos en el modelo.

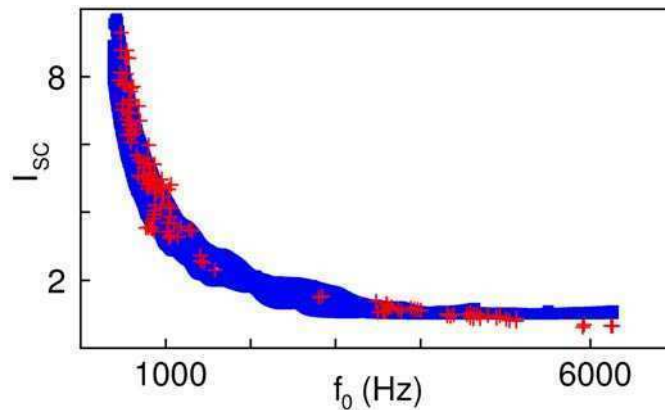


FIGURA 4. Al recorrer una grilla en el espacio de parámetros relacionados con los gestos motores, se sintetizan sonidos integrando el modelo. Los sonidos sintetizados al efectuar este barrido cumplen la relación funcional entre contenido espectral y frecuencia fundamental que se observó en los cantos analizados. En el gráfico se presentan nuevamente los segmentos de canto de la figura 2 y se agregan, representados por cuadrados azules, los puntos correspondientes a los sonidos producidos al efectuar el barrido en el espacio de parámetros  $(p_s, k_1)$ .

Mediante la observación del canto y el modelado de la periferia biomecánica aprendimos que hay una estructura en las propiedades acústicas que viene determinada por la biofísica de la siringe. Durante el aprendizaje del canto, las aves oscilan explorando la relación entre sus gestos motores y las vocalizaciones a las que conducen. Si la interfase entre el gesto motor y la vocalización tiene la capacidad de introducir restricciones a las vocalizaciones que no requieren condiciones particulares en el espacio de parámetros de control motor, cabe preguntarse qué condiciones le imponen estas restricciones a los mecanismos de aprendizaje del canto, entendidos como exploraciones del sistema nervioso central en el espacio de instrucciones motoras.



## Síntesis de canto en tiempo real

El modelo presentado en el capítulo anterior captura mediante características dinámicas de la siringe la diversidad de vocalizaciones posibles para el diamante mandarín. Tiene la potencialidad de producir canto sintético, y si los parámetros de control se pueden relacionar con registros fisiológicos, el canto sintético puede ser generado a partir de gestos motores reales. En este capítulo evaluamos la factibilidad de producir canto sintético en tiempo real a partir de dos registros fisiológicos obtenidos en un ave durante la fonación: la presión sub-siríngea y la actividad del músculo vS.

### 1. Integración en tiempo real mediante DSP

Trabajamos en la implementación de los modelos en tiempo real. Elegimos la plataforma DSP (*Digital Signal Processor*). Los DSP son procesadores dedicados. Del mismo modo que, por ejemplo, una placa de video en una PC tiene incluidos en el *hardware* módulos que realizan en un solo ciclo de reloj una cantidad de instrucciones que son de uso frecuente en el manejo de gráficos, un DSP contiene módulos que realizan operaciones frecuentes en el procesamiento digital de señales. Usando un procesador dedicado, la cantidad de ciclos de operación que se necesitan para cargar las señales digitalizadas en la memoria, procesarlas, y enviarlas a la salida vía un conversor digital/analógico, es menor que la que sería necesaria si se usara una PC. De este modo, aún funcionando a velocidades de reloj más bajas, los DSP permiten implementaciones en tiempo real a velocidades mayores que en las PC. Otra de las ventajas que presentan los DSP es que las implementaciones son portátiles. Los prototipos de la siringe electrónica que presentamos en esta tesis desembocan en la propuesta de construcción de dispositivos bio-prostéticos controlados fisiológicamente. Para ellos, la portabilidad es una característica crítica [SCWM06].

Utilizamos un procesador Texas Instruments TMS320C6713 montado sobre una tarjeta TMS320C6713 DSK. El procesador opera con números en representación de flotante a una frecuencia de 225 MHz [Cha02]. Seleccionamos esta plataforma, y este procesador en particular, porque el chip diseñado específicamente para el procesamiento de señales digitales dispone de una unidad aritmética capaz de realizar, en un único ciclo, operaciones múltiples muy frecuentes en este tipo de tareas (multiplicar/acumular, empujar/tirar en búfer circular, etc [Keh05, Tex01b]). La tarjeta DSK provee al DSP de dos canales de entrada y salida analógicos con frecuencia de muestreo configurable entre 8 y 96 KHz [Tex01a, Tex04].

---

## 2. Simplificando lo simple

La integración numérica de las ecuaciones del modelo es costosa, en términos de cantidad de operaciones, aún tratándose de un sistema de baja dimensión. Para la producción de sonidos de frecuencia relativamente alta (hasta cerca de  $10\text{ KHz}$ ), el tiempo para realizar las operaciones es crítico. Recurriendo a formas comunes en el estudio de sistemas dinámicos, podemos simplificar la computación sin resignar la capacidad descriptiva del sistema.

**2.1. Dinámica del modelo.** En el capítulo 6 introdujimos un modelo de aleteo para la siringe que presenta dos mecanismos dinámicos distintos que originan oscilaciones. La existencia de componentes no lineales en las fuerzas de la restitución de los labios juegan un papel importante en la generación de las oscilaciones de baja frecuencia y alto contenido espectral. En un trabajo anterior [AM08] se demostró que tanto una bifurcación de Hopf como una de nodo-silla en el ciclo límite (SNIC) podría ocurrir. En el primer caso, nacen oscilaciones armónicas con frecuencia finita y la amplitud cero. En el último caso, una par de puntos fijos chocan entre sí, de una manera tal que la variedad inestable de uno pertenece a la variedad estable del otro. Así, una oscilación nace con frecuencia cero, y la amplitud finita. Al barrer un espacio de parámetros vinculado con variables fisiológicas de control motor, se comprobó que el modelo puede reproducir la diversidad de frecuencias y contenidos espectrales observada en el canto del diamante mandarín [SAGM08].

En [AM08] se estudiaron en detalle los distintos tipos de soluciones posibles para las ecuaciones del modelo, poniendo énfasis en las bifurcaciones que dan origen a soluciones oscilatorias. En ese trabajo se realizó un diagrama de bifurcaciones en el espacio de parámetros  $(p_{sub}, k_1)$ , que se muestra en la figura 1a. En la región sombreada, existen tres puntos fijos. Las curvas que definen su frontera corresponden a bifurcaciones nodo-silla, donde un par de puntos fijos, uno estable y uno inestable, chocan entre sí. Las líneas gruesas continuas representan las curvas de bifurcación de Hopf, donde puntos fijos pierden su estabilidad frente a soluciones periódicas. El punto en el que una línea de bifurcación de Hopf se encuentra tangencialmente con una línea de bifurcación nodo-silla es un punto de bifurcación de Takens-Bogdanov: una singularidad lineal con dos autovalores nulos. El conjunto completo de soluciones cualitativamente diferentes que se encuentran en sus proximidades se muestra en la figura 1.

La generación de sílabas con diferentes contenidos espectrales se produce, según el modelo, cuando el ave modula las instrucciones fisiológicas con el fin de seguir caminos en el espacio de parámetros  $(p_{sub}, k_1)$  como los que se indican en la figura. El camino para valores altos de  $k_1$  (regiones 1 y 2) genera sonidos tonales, mientras que un camino para valores de  $k_1$  bajos da lugar a sonidos espectralmente ricos (regiones 5 a 2).

**2.2. Reducción del modelo a la forma normal.** Avanzamos hacia la construcción de un dispositivo que sintetice canto en tiempo real a partir de la integración del modelo. Para hacer esto, es conveniente encontrar un conjunto equivalente de ecuaciones que impliquen operaciones computacionalmente menos costosas. Por equivalente se entiende un conjunto de

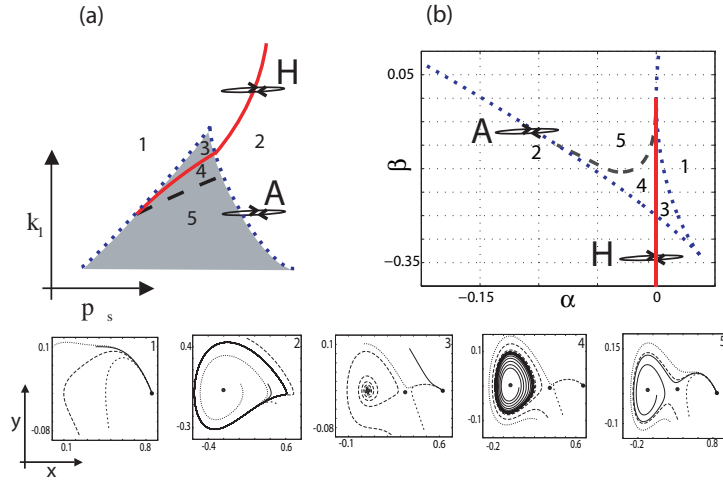


FIGURA 1. Diagrama de bifurcaciones del sistema físico (a) y de la forma normal (b). Para el modelo físico, en el diagrama se muestra el espacio de parámetros presión de los sacos aéreos y tensión  $(p_{sub}, k_1)$ , y para la forma normal el espacio de parámetros  $(\alpha, \beta)$ . La existencia de regiones vecinas como 1 y 2, y las regiones 5 y 2, es la condición matemática para que un modelo sea capaz de generar tanto sonidos tonales y como espectralmente ricos con frecuencias bien definidas. Retratos típicos del espacio de fases  $(x, y)$  para cada región se muestran en los recuadros inferiores correspondientes. En los dos paneles superiores, la línea gruesa continua representa la bifurcación de Hopf, la línea discontinua la bifurcación homoclínica y las líneas de puntos las bifurcaciones nodo-silla. Los caminos A y H representan las diferentes formas de iniciar la oscilación de los labios, cruzando la línea de bifurcación de Hopf (H) y la de SNIC (A), lo que genera sonidos tonales o espectralmente ricos, respectivamente. Extraído de [SAGM10], adaptado de [AM08].

ecuaciones capaz de presentar conjuntos topológicamente equivalentes de soluciones a medida que los parámetros son variados. En particular, es deseable encontrar un sistema de ecuaciones que no involucre un cociente de funciones polinómicas de las variables.

Además de proporcionar una orientación en la tarea de buscar los parámetros de control pertinentes, el análisis dinámico del modelo deja paso a la integración on-line, reduciendo el costo computacional y al mismo tiempo manteniendo la dinámica relevante. Una transformación de coordenadas polinómica alrededor de un punto singular en espacio de parámetros se suele aplicar a los sistemas dinámicos para reducir el sistema a una forma normal, dejando sólo los términos no lineales necesarios para reproducir el conjunto de bifurcaciones en torno a esa singularidad [GH83]. Nuestro modelo descrito por (28, 29) presenta una singularidad de Takens-Bogdanov, donde las líneas de bifurcación de Hopf y SNIC se encuentran tangencialmente. La



reducción a la forma normal alrededor de ese punto elimina muchos términos no lineales y lleva al sistema a una forma en la que el cálculo es mucho más económico [SAGM08].

Empezamos el proceso de simplificación allí donde el sistema presenta una singularidad de *Takens-Bogdanov*. Allí el campo vector puede escribirse, a tercer orden, en su forma normal [AP90]:

$$\begin{cases} \dot{x} &= y \\ \dot{y} &= c_{11}x^2 + c_{12}xy + c_{21}x^3 + c_{22}x^2y. \end{cases} \quad (31)$$

El punto fijo del sistema (28, 29) en la singularidad es

$$\begin{cases} x^* &= \frac{-k_1}{3k_2x_{01}} - \sqrt{\left(\frac{-k_1}{3k_2x_{01}}\right)^2 - \frac{-k_1}{3k_2x_{01}}}, \\ y^* &= 0. \end{cases} \quad (32)$$

Realizando una expansión de Taylor alrededor del punto fijo en los valores de  $p_{sub}^*$   $k_1^*$  donde tanto la condición de nodo-silla y las condiciones de Hopf se satisfacen, obtenemos

$$\begin{cases} \dot{\xi}_x &= \xi_y \\ \dot{\xi}_y &= \left[-3k_2x^* + \frac{p_{sub}(x_{01}-x_{02})}{(x_{01}+x^*)^3}\right] \xi_x - 2 \left[\beta_2x^* + \tau y^* \frac{1+(x_{01}-x_{02})}{(x_{01}+x^*)^3}\right] \xi_x \xi_y + \\ &+ \dots, \end{cases} \quad (33)$$

donde  $\xi_y = y$ ,  $\xi_x = x - x^*$ . Con los valores de los parámetros utilizados en [AM08] para encontrar la bifurcación del modelo físico (ecuación (28)), obtenemos  $c_{11} > 0$  y  $c_{12} < 0$ . Un cambio no lineal de variables nos permite deshacernos del término que va con  $\xi_y^2$ . Este cambio de variables afectará a los coeficientes cúbicos obtenidos en la expansión. Sin embargo, elegimos fijar los términos de saturación fenomenológicos a  $x^3$  y  $x^2y$ . Por lo tanto,  $c_{21}$  y  $c_{22}$  serán negativos. De esta manera, para los parámetros de la bifurcación de *Takens-Bogdanov*, nuestro sistema será equivalente a

$$\begin{cases} \dot{x} &= y, \\ \dot{y} &= x^2 - xy - x^3 - x^2y. \end{cases} \quad (34)$$

Se realizó un despliegue (*unfolding*) de nuestro campo vector como el propuesto por Takens, y el diagrama de bifurcación resultante se muestra en la figura 1b. El diagrama presenta el mismo conjunto de escenarios dinámicos, pero el campo vectorial sólo involucra términos no lineales polinómicos. En otras palabras, se obtuvo el modelo matemático más simple (es decir, con el menor número de términos no lineales) que satisface las condiciones requeridas para la generación de sonidos tanto tonales como espectralmente ricos. Éste presenta un diagrama de bifurcaciones equivalente al del modelo físico (ver figura 1b).

Puesto que este modelo contiene un número mínimo de términos no lineales, se ahorra tiempo de cómputo sin perder las propiedades dinámicas requeridas, manteniendo los parámetros dentro de un rango fisiológico significativo. Introducimos un factor de escala temporal  $\gamma$  para que coincidan la frecuencia de las oscilaciones en las variables del modelo y las fluctuaciones de presión sonora que se asume son causadas por la modulación del flujo de

---

aire en los labios. Con este cambio de variables ( $t \rightarrow t/\gamma$ ,  $y \rightarrow \gamma y$ ) el modelo final queda:

$$\begin{cases} \dot{x} &= y, \\ \dot{y} &= \gamma^2\alpha + \gamma^2\beta x + \gamma^2x^2 - \gamma xy - \gamma^2x^3 - \gamma x^2y, \end{cases} \quad (35)$$

donde  $x$  representa el apartamiento del equilibrio de los labios que oscilan en la siringe,  $\gamma$  es un factor de escala de tiempo, y los parámetros  $\alpha$  y  $\beta$  son funciones de la presión de los sacos aéreos y de la actividad del músculo siríngeo ventral, respectivamente. Estas funciones están definidas por las transformaciones de variables que llevan al modelo físico a la forma normal como se describió anteriormente. Las operaciones básicas que intervienen son: una traslación de los valores de  $(p_s, k_1)$  donde ocurre una bifurcación Takens Bogdanov en modelo físico a  $(\alpha, \beta) = (0, 0)$ , un cambio de escala multiplicativo, y un giro de  $\pi$ .

Se puede generar sonido sintético a partir de esta forma normal del modelo, completándolo con un tracto vocal simple, no interactuante, modelado por un tubo de longitud  $L$ , como el que se presentó en la sección 2.2 del capítulo 6.

### 3. Canto sintético en tiempo real

En esta parte ensayamos la factibilidad de la integración en tiempo real de una frase de canto de diamante mandarín, alimentando a un procesador de señales digitales con los gestos motores fisiológicos registrados previamente en simultáneo con el canto. Primero encontramos caminos en el espacio de parámetros  $(\alpha, \beta)$  que reproduzcan canto comparable al integrar el modelo; luego encontramos transformaciones de los registros fisiológicos (presión sub-siríngea y actividad electromiográfica (EMG) de los músculos siríngeos ventrales  $vS$ ).

**3.1. Caminos de parámetros en la forma normal.** Para ajustar los parámetros del modelo en su forma normal, elegimos dos segmentos del canto cuyo sonograma se muestra en la figura 2 (panel superior). Un segmento corresponde a un sonido tonal de alta frecuencia fundamental, mientras que el otro a un sonido espectralmente rico de baja frecuencia. Encontramos las frecuencias fundamentales (tono) y contenidos espectrales de los dos segmentos. Para ajustar  $\gamma$ , calculamos esos valores para el primer segmento ( $f_0 = 715Hz$ ,  $I_{SC} = 6, 6$ ). Luego integramos las ecuaciones para distintos valores de  $\gamma$  y  $\beta$ . Para cada valor de  $\gamma$ , buscamos el valor de  $\beta$  que determine la frecuencia de oscilación más cercana a la del segmento de canto grabado, y calculamos el índice  $I_{SC}$  del sonido sintético. Se computó, como una función de  $\gamma$ , la distancia entre los valores registrados y los sintetizados  $d_1 = \left( \frac{I_{SC}^{1,s} - I_{SC}^{1,r}}{I_{SC}^{1,s}} \right)^2$ , donde  $I_{SC}^{1,s}$  y  $I_{SC}^{1,r}$  son los índices de contenido espectral de los sonidos sintetizados y de los registrados, respectivamente. El mismo procedimiento se siguió para el segundo segmento ( $f_0 = 5468Hz$ ,  $I_{SC} = 1,0245$ ) para calcular la distancia  $d_2$  definida análogamente. Computamos la cantidad  $d_t = d_1 + d_2$  como una función de  $\gamma$  y encontramos que su mínimo ocurre para  $\gamma = 23500$ . Con el fin de producir caminos en el espacio

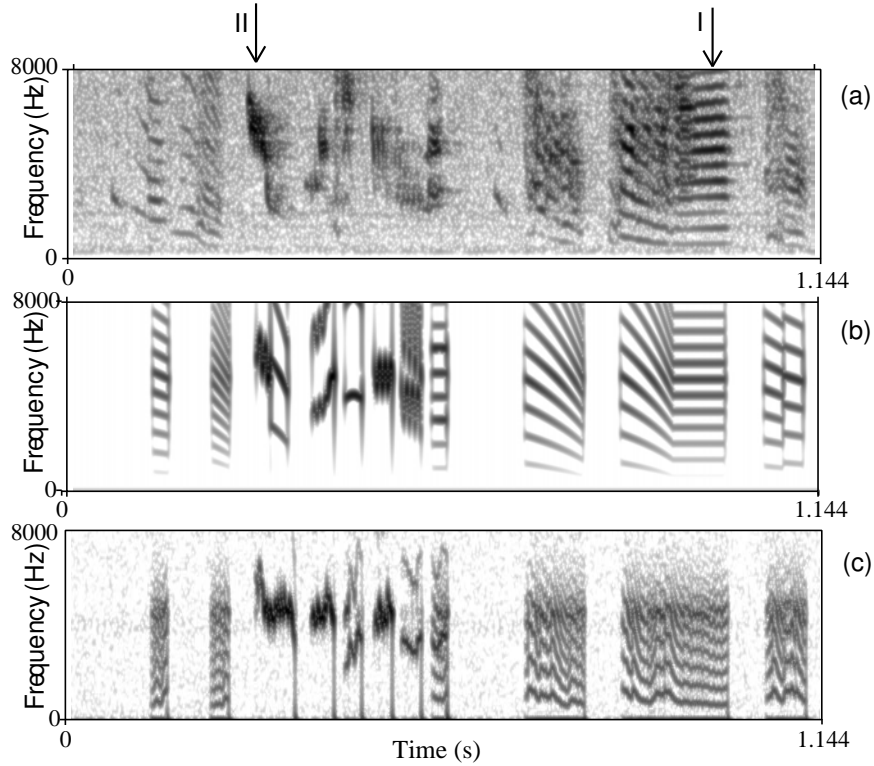


FIGURA 2. Canto natural (panel superior), canto sintético generado por el modelo computacional con parámetros ad-hoc (panel central), salida de la siringe electrónica implementada en un DSP y conducida con registros fisiológicos (panel inferior).  $L = 1,9cm$ ,  $r = -0,9$ . Las flechas indican los segmentos utilizados para el ajuste.

de parámetros que sean capaces de conducir el modelo para generar el sonido sintético similar al registrado, definimos dos regímenes. Suponemos que los sonidos de alta frecuencia (sonidos con frecuencias fundamentales por encima de  $1,8kHz$ ) se producen cuando las oscilaciones de los labios se encienden mediante bifurcaciones de Hopf (un escenario dinámico compatible con oscilaciones tonales [AM08]). Sonidos de baja frecuencia son atribuidos a oscilaciones labiales nacidas en bifurcaciones SNIC [AM08]. En los casos en los que no se observaron discontinuidades espectrales en las sílabas de baja frecuencia, se asume que ambas fuentes se activan en simultáneo y que operan en el mismo régimen dinámico. Construimos dos series temporales  $\beta_{1,2}^{syn}(t)$ , cada una correspondiente a uno de esos regímenes. Cada serie temporal  $\beta_{1,2}^{syn}(t)$  fue construida de tal manera que la frecuencia fundamental del canto sintetizado coincida con la frecuencia fundamental del canto registrado correspondiente a ese instante  $t$ . Cada  $\beta_{1,2}^{syn}(t)$  fue utilizada para conducir el modelo, de forma independiente, para producir dos series de fluctuaciones de presión. Con la suma de esas series se computaron las fluctuaciones de presión a la salida del tracto  $p_o(t)$ , y se produjo canto sintético. El parámetro  $\alpha$  se eligió de modo que la variable  $x$  oscilara al corresponder la fonación.

---

Se optó por un  $\alpha$  que tomara uno de dos valores  $\alpha_{on} = 0,05$  ó  $\alpha_{off} = -0,15$  dependiendo de si había o no fonación, respectivamente. Es pertinente notar que una sílaba de alto contenido espectral que consta de un barrido hacia abajo en frecuencia, que ocurre alrededor de  $t = 0,66s$ , no se sintetiza. Resultados publicados [GC04] sugieren que ese tipo de sonidos puede tener su origen en los efectos no lineales debidos a las interacciones entre las fuentes, no incluidos en este modelo. El canto sintetizado por la conducción de las ecuaciones con  $\beta_{1,2}^{syn}(t)$  se muestra en el panel central de la figura 2.

Una vez identificada una serie temporal de los valores de los parámetros que generan canto sintético con frecuencia fundamental y contenido espectral similares a los del canto grabado, se procedió a: 1) Buscar una transformación que vaya de un registro de actividad (EMG) del músculo  $vS$  de un ave cantando a parámetros dentro del rango calculado anteriormente y 2) implementar este modelo en un procesador digital de señales (DSP).

**3.2. Canto sintético a partir de registros fisiológicos.** El escaneo adecuado de señales EMG de los músculos siríngicos, en el rango de los parámetros dependientes del tiempo calculado previamente, permite sintetizar canto realista. Para registrar la actividad muscular, un par de electrodos fueron implantados en los músculos de la siringe, atravesando el saco aéreo interclavicular. Los electrodos se prepararon a partir de cable aislado de acero inoxidable y fijados al tejido con una microgota de adhesivo tisular. Antes de cerrar el saco aéreo, todos los cables fueron dirigidos hasta el dorso del ave, conectados a un amplificador de instrumentación y a una PC con una placa digitalizadora de señales. Las señales digitalizadas se monitorearon y grabaron utilizando software propio, diseñado y escrito en colaboración con Jorge Alliende, entonces becario pos-doctoral en el Laboratorio de Sistemas Dinámicos. Grabaciones simultáneas de sonido, presión, actividad EMG de los músculos  $vS$  derecho e izquierdo, se muestran en la figura 3. El registro EMG rectificado y suavizado proporciona una medida relativa de la actividad muscular [PB81]. Esto se utilizó para determinar las contribuciones de los músculos a los movimientos responsables de la lenta evolución de la frecuencia fundamental dentro de las sílabas que componen el canto [MGGS03]. Este procesamiento previo de los datos de EMG se realizó con un circuito (ver apéndice A) que rectifica la actividad y la suaviza. El circuito integra la siguiente ecuación:

$$\dot{vS} = -\tau_{EMG}vS + |vS|, \quad (36)$$

con  $\tau_{EMG} = 10ms$  y  $|vS|$  el valor rectificado de la actividad del músculo siríngeo ventral. La salida del circuito se muestra en los paneles A y C de la figura 4. Para seleccionar el conjunto óptimo de señales fisiológicas para conducir el modelo, graficamos la señal  $vS$  suavizada en función del tono. Seleccionamos la actividad  $vS$  del músculo izquierdo para conducir el modelo durante la síntesis de sonidos de baja frecuencia, y la derecha para vocalizaciones de frecuencias fundamentales por encima de  $1,8kHz$ . Esta decisión se tomó después de una inspección de los gráficos de las actividades izquierda/derecha en función del tono. En el panel izquierdo de la figura 5 se muestra, para los sonidos de alta frecuencia, la actividad  $vS$  derecha en función del tono y, para los sonidos de baja frecuencia, la actividad  $vS$

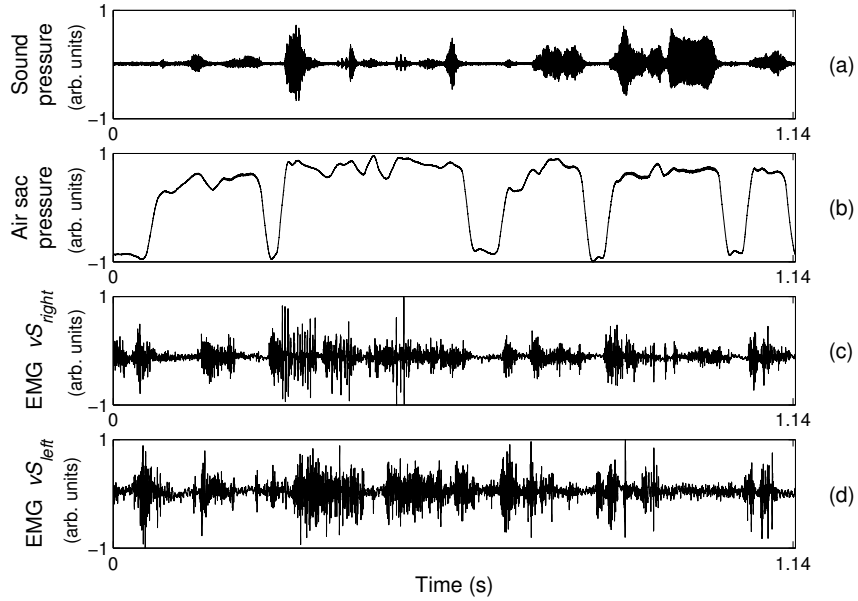


FIGURA 3. Canto grabado de un diamante mandarín (panel superior), canto sintético generado por el modelo computacional con parámetros ad-hoc (panel central), salida de la siringa electrónica implementada en un DSP y conducida con registros fisiológicos (panel inferior).  $L = 1,9cm$ ,  $r = -0,9$ . Las flechas indican los segmentos utilizados para el ajuste.

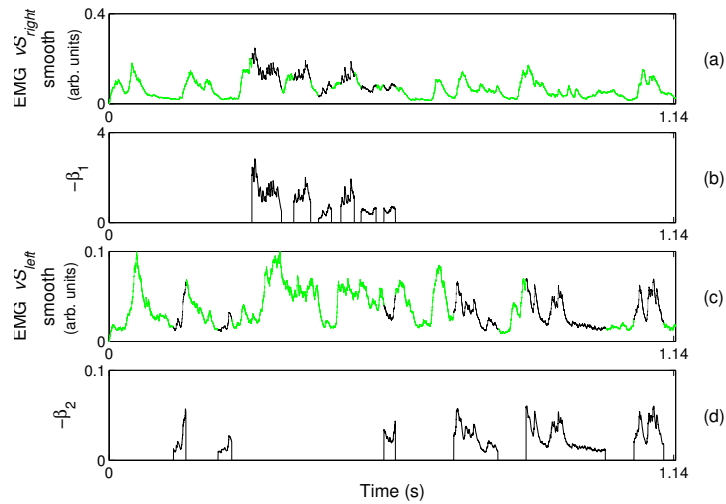


FIGURA 4. Extracción de los parámetros que conducen el modelo a partir de la actividad EMG suavizada de los músculos  $vS$ . Rectificación de la señal de EMG y alisado por filtro pasa-bajos de la actividad  $vS$  derecha (a) e izquierda (c). En ambos paneles, fonación y  $\alpha = \alpha_{on}$  se indican con una línea oscura. No fonación y  $\alpha = \alpha_{off}$  se indican con una línea clara. El parámetro  $\beta^{syn}$  se muestra en los paneles (b) y (d) para los tiempos en los que  $\alpha = \alpha_{on}$ .

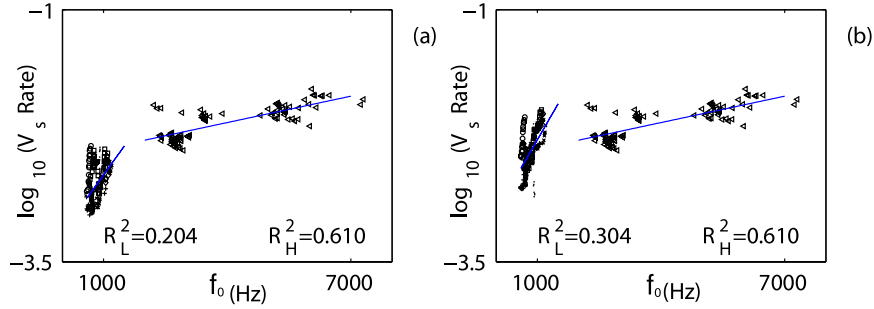


FIGURA 5. Actividad muscular  $vS$  en función del tono,  $f_0$ . En el panel izquierdo, el tono de las sílabas de baja frecuencia ( $f_0 < 1,8kHz$ ) se representa en función de la actividad muscular del lado izquierdo de la siringa, y el de las sílabas de baja frecuencia en función de la actividad  $vS$  derecha. En (b), el tono de todas las sílabas se representa como una función de la actividad muscular derecha.

izquierda en función del tono. En el panel derecho de la figura 5, se muestra la actividad  $vS$  derecha en función del tono para todas las sílabas. En el panel de la derecha, hay niveles de actividad para los que dos frecuencias fundamentales son posibles. Esta ambigüedad sugiere que otros fenómenos podrían estar teniendo lugar, como la actividad de los músculos que no han sido medidos aún, o un efecto no lineal introducido por un acoplamiento entre las fuentes. La única combinación que define una relación biyectiva entre actividad y frecuencia tiene lugar cuando la actividad izquierda se utiliza para sonidos de baja frecuencia y la actividad derecha para los sonidos de alta frecuencia, por lo que es la mejor opción para conducir el modelo.

Para sintetizar canto mediante la conducción del modelo con los registros fisiológicos, tenemos que encontrar una transformación de las actividades registradas  $EMG$  a los parámetros del modelo  $(\alpha, \beta)$  que se calcularon con anterioridad. Como antes,  $\alpha$  se toma como uno de los dos valores que colocan el modelo dentro o fuera de un régimen oscilatorio. Se midió la presión del saco aéreo mediante la inserción de una cánula en el saco torácico, con el otro extremo conectado a un transductor de presión (ver [MGGS03]). Las condiciones para  $\alpha = \alpha_{on}$  fueron que la presión medida supere un valor umbral y que el ave estuviera fonando. Los parámetros  $\beta$  son los responsables de la modulación de frecuencia y el contenido espectral de los sonidos sintetizados. Buscamos transformaciones de las actividades musculares que lleven a las series  $\beta^{syn}$  construidas artificialmente para que el modelo reproduzca sonido similar al canto grabado. Propusimos

$$\beta_i^{fit} = a_{0,i} + a_{1,i}vS_i(t) + a_{2,i}vS_i^2(t) \quad (37)$$

para relacionar la actividad muscular ventral con los parámetros en la forma normal del modelo. Elegimos  $a_{j,i}$  de modo que  $\chi^2 = \chi^2(\beta_i^{fit} - \beta_i^{syn})$  sea mínimo, con  $i = 1, 2$  correspondientes a regímenes de frecuencia alta o de baja para  $\beta$ , y derecha o la izquierda para los registros de actividad de  $vS$ , respectivamente. Estos parámetros fueron ajustados con tres pares de valores  $\beta_i^{fit}$

---

y  $\beta_i^{syn}$ , a tiempos alrededor del segmento II de la figura 2, para alta frecuencia. Para baja frecuencia se procedió análogamente a tiempos alrededor del segmento I. En nuestro ejemplo, los valores obtenidos son  $(a_{0,1}, a_{1,1}, a_{2,1}) = (-6697, 152,65, 0,0848)$  y  $(a_{0,2}, a_{1,2}, a_{2,2}) = (0, -17,79, 0,0016)$ . Hemos utilizado sólo unos pocos puntos en los dos segmentos para realizar el ajuste, y luego reproducido el canto entero. En ese sentido, el modelo muestra tener un poder predictivo. Los parámetros  $\beta_i^{fit}$  a tiempos en los que  $\alpha = \alpha_{on}$  se muestran en los paneles b y d de la figura 4. Con este ajuste, somos capaces de controlar fisiológicamente un dispositivo que integre las ecuaciones del modelo.

**3.3. Integración en tiempo real.** Implementamos nuestro modelo en un DSP montado en una tarjeta Texas Instruments TMS320C6713 DSK. La placa se compone de una serie de canales de entrada/salida analógicos, memoria, y un DSP que opera a 225 MHz. Se escribió un programa en C que hace uso de las capacidades de muestreo de la placa para digitalizar los datos correspondientes a las grabaciones de la actividad fisiológica, realizar una integración numérica del modelo, y convertir la salida en una señal analógica (el canto sintético).

Con el fin de realizar una integración en tiempo real del modelo del órgano vocal, utilizamos estas transformaciones para definir los parámetros para alimentar al DSP. La placa digitalizó los datos a una frecuencia de muestreo de  $48kHz$ . Dentro de cada paso de muestreo, el procesador integró las ecuaciones que describen la dinámica de los labios. El método de Euler fue elegido debido a su bajo coste computacional, lo que permitió realizar 18 iteraciones en el tiempo entre muestras. Dos integraciones independientes del modelo se realizaron. Cada variable  $x_{1,2}(t)$  resuelta numéricamente se convirtió en una señal analógica por la DSP en la misma frecuencia a la que la entrada se muestreaba. Las soluciones se utilizaron para generar una señal de sonido, cuyo sonograma se muestra en la figura 2 (panel inferior).

Construimos así un dispositivo electrónico que es capaz de traducir las instrucciones fisiológicas en canto en tiempo real. Lo hicimos mediante la programación de un procesador digital de señales (DSP) para que el dispositivo pueda leer los datos experimentales, integrar las ecuaciones dinámicas de un modelo físico para el problema, y el uso de estas soluciones para sintetizar el canto. El uso de la tecnología digital es una mejora cualitativa respecto de las anteriores soluciones de voz bio-miméticas [ZMP<sup>+</sup>05]: el control de los parámetros del sistema, ya sea para mejorar el modelo o como una fuente realimentación sensorial alterada, es flexible y preciso.

## Interfaz cerebro-máquina para diamante mandarín

Los dispositivos de Interfaz Cerebro Máquina (BMI) decodifican las instrucciones motoras a partir de registros de algún tipo de actividad neurofisiológica y conducen efectores bio-miméticos para producir retroalimentación sensorial sintética. Muchas aplicaciones logran una alta precisión en un número limitado de tareas mediante la aplicación de métodos estadísticos a estos datos para extraer las características correspondientes a determinadas instrucciones motoras [SCWM06].

Construimos un órgano vocal aviar bio-prostético. El dispositivo se basa en un modelo matemático de baja dimensión que da cuenta de la dinámica del órgano vocal del diamante mandarín y relaciona caminos suaves en un espacio de parámetros fisiológicamente significativo con secuencias de vocalizaciones complejas.

Los dos gestos fisiológicos motores (presión sub-siríngea y actividad del músculo siríngeo ventral), se reconstruyen a partir del canto del ave, y el modelo se implementa en un procesador digital de señales para producir canto sintético al ser controlado a través del gesto de presión por un pájaro comportándose libremente.

Esto ejemplifica la plausibilidad de un tipo de interfaz entre el cerebro y un comportamiento complejo sintético. En este tipo de dispositivos, la comprensión de la biomecánica de la periferia es clave para la identificación de un conjunto de señales fisiológicas de baja dimensión que codifiquen las instrucciones motoras; por lo tanto, permite la implementación en tiempo real a un bajo costo computacional.

### 1. BMI con modelo de la biomecánica

**1.1. BMI como lo conocemos.** La estrategia habitual de BCI y BMI consiste en descifrar los comandos motores a partir de registros de actividad fisiológica en el cerebro y utilizar esta actividad para controlar dispositivos bio-miméticos [SCWM06, CWR<sup>+</sup>08, CLC<sup>+</sup>03]. En [CLC<sup>+</sup>03], por ejemplo, registros de electrodo múltiple de decenas a cientos de neuronas en diferentes áreas de la corteza de los primates se utilizan para accionar un brazo robótico. En un trabajo reciente, Cichocki *et. al.* discuten las perspectivas de la utilización de registros electroencefalográficos (EEG) para generar soluciones BCI no invasivas [CWR<sup>+</sup>08].

En estos ejemplos (como en muchas otras implementaciones de BCI), el problema crucial es la clasificación de las características de la gran cantidad de datos que corresponden a un conjunto determinado de tareas motoras. La extracción de estas características (*feature extraction*) se realiza mediante diferentes técnicas que incluyen la descomposición lineal en una diversidad



---

de espacios vectoriales, así como algoritmos de aprendizaje automático (*machine learning*) [SCWM06, PS00, CWR<sup>+</sup>08]. De esta manera, se logra un control preciso de efectores bio-miméticos para un número finito de tareas específicas, como mover un cursor en una pantalla o tomar un objeto con un brazo robótico. Debido a que la extracción de características y la clasificación se realiza sobre bases estadísticas en un espacio de parámetros de alta dimensión, se requiere de cálculo pesado. Esto impone una restricción a la portabilidad de estas soluciones y a la complejidad de las tareas a realizar.

**1.2. El aporte del modelo.** En el capítulo 6 presentamos un modelo basado en el mecanismo de aleteo propuesto por Titze para las oscilaciones en las cuerdas vocales humanas [Tit88], y lo utilizamos luego para sintetizar el canto del diamante mandarín en tiempo real [SAGM10, PAA<sup>+</sup>11]. Este modelo captura la dinámica no lineal de las membranas que oscilan para producir sonido, de modo tal que una variedad de vocalizaciones complejas es generada mediante un ajuste de los parámetros relacionados a gestos motores fisiológicamente observables suscitado por el ave.

Parte del atractivo de contar con este modelo está en la posibilidad de aplicarlo a la construcción de un dispositivo bio-protésico. En este escenario el cálculo es relativamente económico, debido a la baja dimensión del modelo matemático. Además, la descripción física de los efectores periféricos condujo a la identificación de un conjunto de parámetros suavemente variables que determinan el comportamiento [SAGM10]. Mediante el registro de la actividad fisiológica relacionada con los parámetros y su paso a un dispositivo que resuelve las ecuaciones del modelo *on-line*, se puede emular comportamiento vocal por una prótesis controlada por un sujeto a través de sus instrucciones motoras.

En este capítulo presentamos un dispositivo que es conducido por un diamante mandarín comportándose libremente para producir vocalizaciones sintéticas realistas *on-line*. El dispositivo se basa en la integración en tiempo real del modelo matemático del órgano vocal en un procesador de señales digitales (DSP). Es controlado por el gesto de presión del saco aéreo torácico del ave, que es transducido, digitalizado y enviado al DSP para proporcionar al modelo el camino apropiado en el espacio de parámetros.

## 2. Un tracto vocal superior más completo

Los labios que oscilan en la siringe modulan el flujo de aire para la producción de sonido, que se ve modificado a su paso por el tracto vocal. Características acústicas relevantes, tales como el contenido espectral de las vocalizaciones, están determinadas por la geometría del tracto vocal [Tit00, SAGM08, PAA<sup>+</sup>11]. La tráquea, la cavidad oro-faríngea-esofágica (OEC) y el pico son elementos importantes para la modulación del canto en el tracto vocal, como fuera notado por Nowicki [Now87] y Fletcher *et. al.* [RSFB06, FRS06].

Con el fin de producir vocalizaciones sintéticas más realistas que en el capítulo anterior, introducimos un modelo del tracto vocal como un sistema dinámico, que incluye un tubo que aproxima la tráquea y un resonador de Helmholtz para representar a la OEC. El sonido producido en la siringe

entra en este sistema, y somos capaces de computar la el sonido radiado a la atmósfera.

La tráquea, por su efecto y por su fisiología, se aproxima por un tubo de longitud  $L$  que está cerrado al extremo siríngeo y abierto en el glotal. La parte transmitida de la fluctuación de presión  $p_o(t) = (1 - r)p_i(t - L/v_s)$  fuerza el aire en la glottis, que es aproximada por el cuello del resonador Helmholtz que representa la OEC. La masa de aire en la glottis, forzada hacia adentro de la cavidad, está sujeta a una fuerza de restitución ejercida por la masa de aire mayor en ella.

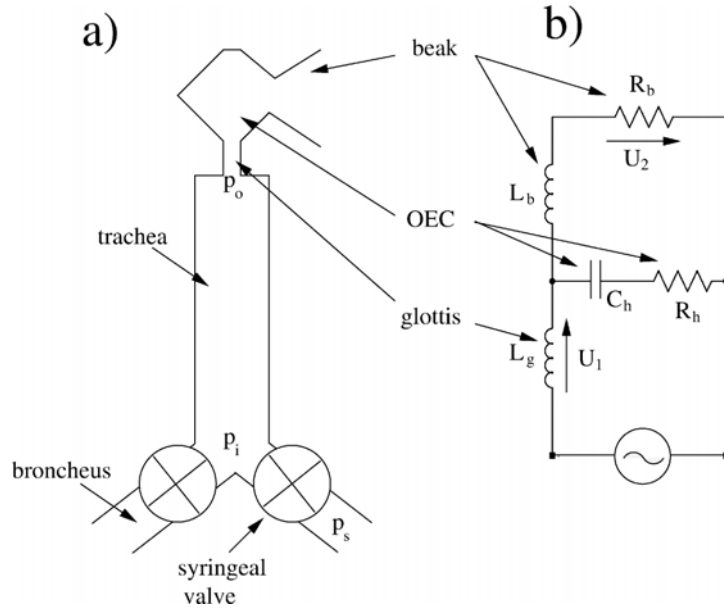


FIGURA 1. Diagrama de nuestro modelo para el aparato vocal del diamante mandarín (a). El sonido es producido en las válvulas siríngeas, y luego es filtrado en su paso a través del tracto vocal. La presión a la entrada del tracto es  $p_i$  y  $p_o$  es la presión a la salida de la tráquea. Para computar el modelo escribimos el circuito eléctrico equivalente de la parte post-traqueal del tracto vocal (b).

En acústica, es común escribir un modelo computacional análogo electrónico para describir un sistema de filtros. La presión acústica es representada por un potencial eléctrico y el caudal por la corriente eléctrica [CFKS82]. En este marco, las constricciones cortas son inductores, y las cavidades (más pequeñas que las longitudes de onda) son bien representados por capacitores. Las ecuaciones del circuito equivalente de la parte posterior a la tráquea del tracto vocal, que se muestra en la figura 1 b, quedan:

$$\begin{cases} \frac{di_1}{dt} = \Omega_1, \\ \frac{d\Omega_1}{dt} = -\frac{1}{L_g C_h} i_1 - R_h \left( \frac{1}{L_b} + \frac{1}{L_g} \right) \Omega_1 + \\ \quad + i_3 \left( \frac{1}{L_g C_h} - \frac{R_b R_h}{L_b L_g} \right) + \frac{1}{L_g} \frac{dV_{ext}}{dt} + \frac{R_h}{L_g L_b} V_{ext}, \\ \frac{di_3}{dt} = -\frac{L_g}{L_b} \Omega_1 - \frac{R_b}{L_b} i_3 + \frac{1}{L_b} V_{ext}, \end{cases} \quad (38)$$

---

donde los componentes eléctricos se relacionan con los parámetros geométricos de los elementos acústicos. Estas relaciones, que son estándar en acústica, se pueden encontrar en [PAA<sup>+</sup>11, CFKS82].

**2.1. BMI para un diamante mandarín.** El modelo matemático para el órgano vocal, la reducción del sistema que rige la dinámica de la fuente de sonido, y la identificación de los parámetros pertinentes que dan cuenta de su control motor, todos ellos abren paso a la construcción de un dispositivo bio-protésico. Los parámetros que determinan las propiedades acústicas de las vocalizaciones en la forma normal (35) son fisiológicamente significativos y el conjunto de ecuaciones diferenciales es suficientemente fácil de calcular en una plataforma portátil, como un procesador de señales digitales. Mediante el ajuste de los parámetros y la integración del sistema en tiempo real, puede producirse canto sintético en un dispositivo controlado por las instrucciones motoras generadas por un pájaro comportándose libremente.

En muchas soluciones de bio-prótesis, los registros fisiológicos utilizados para accionar el dispositivo están degradados respecto de los registrados en los sujetos intactos [SCWM06]. Una de las aplicaciones de este dispositivo es la realización de experimentos de realimentación auditiva alterada [BD00, CG04, SB06]. En los experimentos habilitados por este dispositivo, la realimentación auditiva propia de las aves puede ser sustituida por canto sintetizado *on-line*. Ya que la realimentación sintética es producida por la integración del modelo cuando es conducido con gestos motores fisiológicos suscitados por un pájaro comportándose libremente, es posible introducir alteraciones en la realimentación que sean compatibles con alteraciones en los gestos motores destinados a producirlas. En estos experimentos el órgano protésico vocal es conducido por un ave que ha sido silenciada mediante la inserción de una cánula a través de su saco aéreo inter-clavicular. La fonación se impide porque el flujo de aire se desvía por fuera de la siringe. Como efecto secundario, el patrón de presión registrado en el ave enmudecida difiere del registrado en el animal intacto.

El dispositivo reconstruye el gesto motor intentado a partir de este gesto de presión degradado para activar la de integración del modelo. Cuando el patrón de presión correspondiente a las sílabas que comprenden un motivo son identificadas, el modelo matemático del órgano vocal se integra con los caminos adecuados en el espacio de parámetros para producir la síntesis correspondiente.

### 3. Implementación

La siringe electrónica es capaz de reproducir canto sintético en tiempo real cuando es conducida por la presión del saco aéreo de un ave enmudecida comportándose libremente. El gesto de presión es registrado, junto con la canción del pájaro, con el fin de ajustar los parámetros del modelo. El animal es enmudecido a través de un desvío del flujo de aire por fuera de la siringe y su gesto de presión se digitaliza y se alimenta al DSP, donde se implementa el modelo. Un algoritmo que reconstruye el gesto de presión del ave intacta es implementado, para activar la integración del modelo cuando el gesto correspondiente a la primera sílaba de un motivo es detectado y

---

apagarla cuando corresponde. A continuación se presenta una descripción del procedimiento.

**3.1. Registro del gesto de la presión de saco aéreo.** El pájaro es canulado. Una cánula (*Silastic Laboratory Tubing*, 0,76 mm de diámetro interno, 1,65 mm de diámetro externo) se inserta a través de una incisión en el saco aéreo torácico, a través de la piel, en la zona abdominal del ave <sup>1</sup>. El otro extremo de la cánula es conectado a un transductor piezoeléctrico *Fujikura FPMC-07PGR*, amplificado y digitalizado, simultáneamente con el sonido producido por las aves (grabado con un micrófono *Takstar SGC568*).

La alimentación eléctrica del transductor, la amplificación de la señal y el control de su nivel de continua son realizados mediante electrónica que construimos con ese propósito (ver apéndice B). El diseño del circuito que realiza esa tarea fue hecho de modo que se puedan controlar independientemente el nivel de tensión continua (*offset*) y la amplitud de la señal. Se requiere esta flexibilidad, además de estabilidad una vez efectuada la calibración, para que el gesto de presión ocupe de modo apropiado el rango del instrumento con el que se digitalizará. En esta etapa, el conjunto de datos se introduce en una PC mediante un conversor analógico-digital *National Instruments NI-6259*, que muestrea a  $24\text{KHz}$  y con una resolución de  $16\text{ bit}$  un rango de  $\pm 2,5\text{ Volt}$ .

El pájaro se coloca en una caja acústica (construida en colaboración con los estudiantes de la Lic. en Ciencias Biológicas Rodrigo Alonso y Nicolás Adreani, y los estudiantes de Doctorado Yonatan Sanz Perl y Matías Goldin), y el sonido y la presión se registran mientras canta espontáneamente. En los paneles superiores de la figura 4 están ilustrados el sonograma de una frase de canto y su correspondiente gesto de presión.

**3.2. Reconstrucción del gesto motor.** Estos datos son utilizados para ajustar los parámetros fisiológicamente significativos del modelo del órgano vocal en su forma normal (35). Las características de los sonidos producidos por la integración de la forma normal del modelo están determinadas por caminos en el espacio de parámetros  $(\alpha, \beta)$ , que son transformaciones de la presión sub-siríngea y de la actividad del músculo siríngeo ventral [SAGM10, PAA<sup>+</sup>11]. Con el fin de producir vocalizaciones sintéticas, el gesto motor es reconstruido a partir del canto grabado, en busca de las excursiones llevan  $(\alpha(t), \beta(t))$  que a las soluciones de iguales propiedades acústicas.

En el capítulo anterior y en un trabajo previo [SAGM10] encontramos caminos en el espacio de parámetros que condujeron a síntesis comprobables con el canto real, utilizando el modelo de aleteo extendido para la fuente y un tracto vocal superior sencillo, que contenía apenas un tubo de longitud  $L$  para reproducir el efecto de filtrado a través de la tráquea. En esta parte introducimos un modelo más detallado del tracto vocal, persiguiendo el fin de reproducir canto más realista. Con este objetivo revisamos también el procedimiento para encontrar los parámetros para conducir el modelo.

---

<sup>1</sup>Las intervenciones quirúrgicas se hicieron bajo la supervisión de la Veterinaria María de los Ángeles Suárez, siguiendo los protocolos de cuidado animal del *National Institute of Health* de Estados Unidos.

---

Dos características importantes que se busca igualar son la riqueza espectral y la frecuencia fundamental de las vocalizaciones. Una forma de cuantificar la riqueza espectral es calcular el índice de contenido espectral ( $I_{SC}$ ) de un segmento de sonido [SAGM08].

El parámetro  $\gamma$  en (35) determina la escala temporal. Con este y los parámetros de la descripción del tracto vocal fijos, una frecuencia fundamental y contenido espectral ( $f_0, I_{SCI}$ ) corresponden a cada punto en el espacio  $(\alpha, \beta)$ . La dependencia de las propiedades del sonido de los parámetros se entiende en términos de la distancia, en el espacio de parámetros, desde las líneas de bifurcación del modelo. Cuanto más lejos está un punto de las bifurcaciones de Hopf y SNIC, más altos son su índice de contenido espectral y frecuencia fundamental correspondientes [SAGM08, AM08].

El valor óptimo del factor de escala  $\gamma = 24000$  se encontró minimizando la distancia acumulada en el espacio de propiedades acústicas ( $f_0, I_{SC}$ ) entre 28 segmentos de canto grabado (20 ms de largo) y sus síntesis objetivo. Para cada valor de  $\gamma$  dentro de un rango, se sintetizaron sonidos correspondientes a una grilla de puntos en el espacio  $(\alpha, \beta)$ , y su frecuencia fundamental e índice de contenido espectral fueron calculados. Para cada segmento grabado, se encontró el sonido sintético que minimiza la distancia en el espacio ( $f_0, I_{SC}$ ). La suma de estas distancias mínimas se realizó sobre todos los segmentos grabados, para cada  $\gamma$  en el intervalo. El valor de  $\gamma$  para el que la distancia total es la más pequeña es el óptimo.

Estableciendo también los parámetros del tracto para que la resonancia de la OEC se encuentre cerca de  $4,5\text{ KHz}$  y el tubo que representa la tráquea tenga de  $3,5\text{ cm}$  de longitud, se reconstruye el gesto motor que sintetiza un motivo del canto. Para ello, el canto grabado (muestreado a  $24\text{ KHz}$ ) se descompuso en una secuencia de segmentos de  $20\text{ ms}$ . Para cada segmento, se calcularon la frecuencia fundamental y el índice  $I_{SC}$ . Una búsqueda en el espacio de parámetros  $(\alpha, \beta)$  se llevó a cabo hasta que se encontraron los valores que permiten sintetizar los sonidos con las características acústicas más similares a las de las grabaciones para el rango de parámetros disponible. Para ello, se computó la frecuencia fundamental para una grilla en el espacio de parámetros, y el conjunto de pares  $(\alpha, \beta)$  que produce los sonidos sintéticos más de frecuencia fundamental más cercana a la de la grabación fue seleccionado. Dentro de ese conjunto, el valor de  $(\alpha, \beta)$  que minimiza la distancia en el valor de  $I_{SC}$  fue elegido.

Por este medio y tras un suavizado para interpolar los valores de los parámetros dentro de cada segmento, se obtiene una tabla de valores de los gestos motores reconstruidos para cada segmento del motivo de canto. Cuando el modelo es conducido por ellos, se produce canto sintético similar al que se registró. Una serie de gestos motores para la forma normal construidos de este modo se muestran en la figura 2, junto con el canto sintetizado a partir de ellos y el canto grabado.

En [PAA<sup>+</sup>11], utilizamos los valores de  $\alpha$  y  $\beta$  reconstruidos de esta manera para estimar los valores de la presión del saco aéreo y la actividad del músculo siríngeo ventral ( $vS$ ) a través de un cambio de signo y la multiplicación por un factor de escala [SAGM10]. Cuando los gestos reconstruidos fueron comparados contra los valores reales, grabados, el acuerdo fue fuerte.

---

**3.3. Enmudecimiento.** El animal es enmudecido mediante la inserción de una cánula (*Braintree Scientific MRE 05*, 1,02 mm de diámetro interno, 1,27 mm de diámetro externo) a través del saco aéreo interclavicular. El flujo de aire es desviado entonces de la válvula siríngica, y se previene que ocurra la fonación. El enmudecimiento logrado de este modo tiene como efecto la degradación del gesto de presión registrado. Por otro lado, es reversible. Un diamante mandarín recupera el canto, al menos parcialmente, luego de 4 – 7 días de la intervención.

La presión de saco aéreo del ave enmudecida se registra mientras ésta intenta producir canto, mediante el gesto correspondiente a llamados y a frases.

**3.4. Control de la integración con el gesto de presión degradado.** Este registro se utiliza para ajustar el algoritmo que calcula la correlación del gesto presión enmudecido con el gesto de presión intacto para reconocer la primera sílaba del canto. Se selecciona un segmento de 10 – 150 ms del gesto de presión del ave intacta que contiene la última parte de la nota introductoria precedente y la primera parte de la primera sílaba del motivo. El gesto de presión alterado  $p_m(t)$  es recorrido, y a cada punto se le mide la correlación con el segmento muestreado en el canto intacto. Para obtener la correlación, se toma a partir de cada punto una ventana de la misma longitud  $\Delta$  que el segmento patrón tomado del gesto  $p_u(t)$  del ave intacta y se calcula

$$\rho(t) = \frac{\int_0^\Delta (p_u(t') - \bar{p}_u) (p_m(t + t') - \bar{p}_m) dt'}{\int_0^\Delta (p_u(t') - \bar{p}_u)^2 dt'}, \quad (39)$$

con  $\bar{z}$  denotando valor medio de  $z$  en la ventana de longitud  $\Delta$ .

Hacemos una exploración *off-line* para seleccionar el segmento del gesto intacto que mejor permita la discriminación del principio del motivo, reduciendo al mínimo el retraso en la detección. El segmento patrón debe incluir la menor cantidad posible de muestras de la primera parte del motivo, porque ese tiempo requerido para la detección, al emplearse sobre un flujo de datos que se procesan *on-line*, será tiempo del motivo que no se podrá reproducir al momento que el ave intenta hacerlo mediante su gesto motor. En la exploración, buscamos el segmento que cumpla esta condición y que pueda ser diferenciado suficientemente bien de los llamados y las notas introductorias. Hacemos una estadística sobre no menos de 50 llamados y notas introductorias y 10 motivos, y buscamos que la correlación media para el motivo supere a la correlación media para los demás casos en al menos 1,5 veces la desviación estándar de esta última.

El criterio de elección de esos números para la estadística tiene que ver con los tiempos del procedimiento y la tasa de éxito. La cantidad de llamados y notas introductorias utilizadas para esta calibración es la disponible luego de un día de registro de datos (el día posterior al enmudecimiento). El umbral de correlaciones propuesto permite tener éxito en la detección del motivo en una buena parte de los casos, evitando casi por completo los falsos positivos (falsa detección del motivo ante un llamado o una nota introductoria).

En el panel C de la figura 2 se muestra la correlación calculada de este modo a lo largo de un segmento de gesto de presión del ave enmudecida.

El segmento incluye llamados y dos frases que comienzan cada uno con un número distinto de notas introductorias. Cuando se supera el umbral de correlación, se enciende la integración. En el panel D de la misma figura, un segmento del gesto de presión se muestra junto a la síntesis de canto al detectarse con éxito el inicio de cada motivo. Notar que en este caso, el segmento patrón usado para la detección del inicio del motivo incluyó solo sus primeros 15 ms.

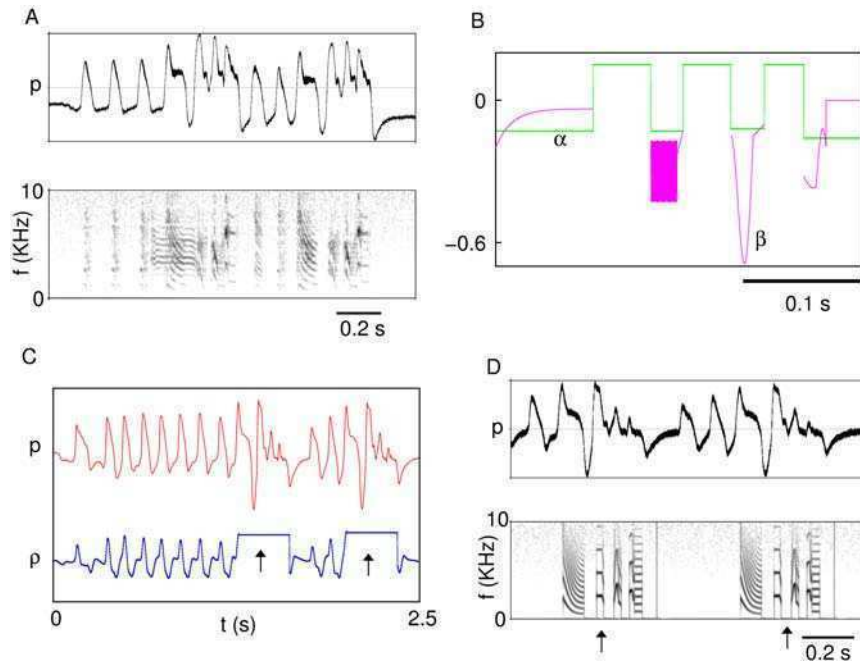


FIGURA 2. Ilustración del procedimiento de ajuste y calibración de la siringe electrónica. El gesto de presión del saco aéreo torácico se registra junto con el canto; un gesto de presión y el sonograma de la frase correspondiente están en la parte A de la figura. A partir del canto, se construyen las series temporales  $(\alpha(t), \beta(t))$  que originan síntesis de las sílabas correspondientes al motivo (parte B). Luego de enmudecer el ave y registrar el gesto de presión que emite al intentar producir canto, se afina el algoritmo que detecta, a partir de este gesto, la ocurrencia de las sílabas que comprenden el motivo. En la parte C de la figura está el gesto de presión degradado y el resultado de calcular la correlación. Los segmentos indicados por flechas indican esta detección. Durante ese tiempo, se sintetiza canto utilizando los parámetros encontrados previamente, como se muestra en la parte D, donde están el gesto de presión del pájaro enmudecido y la salida del algoritmo de detección/integración.

También sintonizamos un algoritmo que detecta la interrupción del canto, comparando las propiedades del gesto de presión enmudecido con umbrales de valor absoluto y variación del gesto intacto durante el motivo ulterior.

---

El canto del diamante mandarín está altamente estereotipado. Una vez que un motivo comienza, se sigue una secuencia de sílabas que es ubicua, y una frase de canto se basa en la repetición del motivo, acaso con notas introductorias intercaladas entre una y otra repetición [ZM04]. Para el control de la siringe electrónica con el gesto de presión, basta saber cuándo comienza un motivo y cuando se interrumpe. Mientras no se interrumpa, la vocalización adecuada se sintetiza integrando el modelo a partir la secuencia de gestos motores correspondientes. Determinamos que esa secuencia es recorrida de a ventanas de  $10 - 20\text{ ms}$ , preguntándose algorítmicamente al final de cada ventana si corresponde integrar el siguiente segmento, o cesar la fonación. El cese de la fonación se produce cuando el gesto de presión se mantiene por debajo de un determinado umbral luego de haberse registrado una pendiente negativa.

El método utilizado para activar la integración del modelo introduce un retardo, porque un segmento de la primer sílaba del motivo se utiliza para detectar su comienzo en el gesto de presión del ave enmudecida. La longitud de este segmento determina el tiempo que tarda la síntesis en comenzar después de que el pájaro ha intentado producir la sílaba. La síntesis, que comienza luego de la detección, se realiza a partir del primer instante de canto que corresponda luego de ese segmento. De esta manera, no se introduce un retardo en la síntesis.

**3.5. El dispositivo.** El modelo con los caminos de parámetros construidos ( $\alpha(t)$ ,  $\beta(t)$ ) y el algoritmo de detección son programados en una plataforma de desarrollo DSP (*Texas Instruments DSK6713*). Esta plataforma contiene una memoria, un conjunto de conversores AD/DA para entrada y salida y DSP un punto flotante funcionando a 225 MHz.

El pájaro enmudecido se coloca en la caja acústica y su gesto de presión es registrado y alimentado al DSP, donde el algoritmo de detección corre en línea. El gesto de presión es muestreado a  $500\text{ Hz}$  y con una resolución de  $12\text{ bit}$  en un rango de  $[0 : +2,5]\text{ Volt}$  por una tarjeta adicional (*DSK6XXXHPI Daughtercard, Educational DSP*), que se comunica con la tarjeta de desarrollo DSK mediante un puerto rápido. Este flujo de datos pasa a través del algoritmo de detección.

Cuando el comienzo de un motivo se infiere a partir del gesto de presión del ave enmudecida, se integra numéricamente el modelo, llamando los valores correspondientes de los parámetros, que fueron previamente guardados en la memoria de la placa DSK. La integración numérica del modelo se lleva a cabo mediante el método de Euler con pasos de  $0,83\ \mu\text{s}$ , de manera que la salida pueda ser generada a  $24\text{ KHz}$ . Esta salida, que corresponde al sonido producido por la siringe, es usada para calcular la presión en la salida de la tráquea y enviada a otro DSP que computa las ecuaciones que dan cuenta de la dinámica de la parte del tracto vocal superior a la tráquea (la glotis, la OEC, y el pico). El sonido sintético producido por este medio se convierte en una señal analógica, se amplifica y se reproduce a través de un parlante, situado en la caja acústica, distante en alrededor de  $20 - 30\text{ cm}$  del ave.



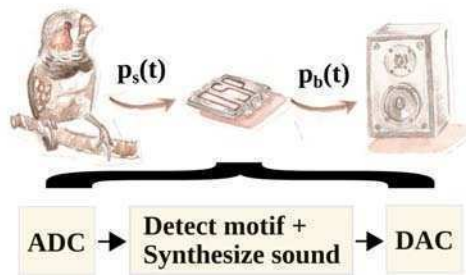


FIGURA 3. Esquema de la configuración experimental en la que el ave enmudecida conduce el órgano vocal electrónico. El gesto de presión del pájaro es registrado simultáneamente con el canto. Luego, los parámetros que conducen la forma normal para producir canto sintético son construidos, y el ave es enmudecida. El pájaro es luego conectado con la siringe electrónica mediante la presión de su saco aéreo torácico, que es digitalizada y enviada al DSP. En el DSP, un algoritmo detecta el comienzo de la primer sílaba del motivo en el gesto de presión degradado del ave enmudecida. Sobre la detección, se integra el modelo *on-line* mientras el intento del pájaro de continuar con el motivo sea inferido a partir del gesto motor. El sonido computado a la salida del pico es convertido en una señal analógica y reproducido a través de un parlante situado a 20-30 cm del pájaro.

El sonido registrado en la caja a través del micrófono, la salida directa analógica del sistema DSP y el gesto de presión alterado se digitalizan y se registran a  $24\text{ KHz}$  en una PC.

**3.6. El pájaro mudo que canta.** El dispositivo tiene éxito en sintetizar canto en tiempo real cuando es conducido por un pájaro enmudecido comportándose libremente a través de su gesto de presión. A partir del gesto motor alterado, el algoritmo infiere el segmento del motivo evocado por el ave y computa el modelo para producir las vocalizaciones. Un ejemplo se ilustra en la figura 4.

Los paneles superiores de la figura muestran la presión sub-siríngea registrada y el sonograma de un segmento de un canto con su nota introductoria precedente. Una frase de canto de este ave (B06) se compone de una serie de notas introductorias (O) y la repetición de un motivo sencillo que contiene dos sílabas (A y B), indicadas por diferentes colores y opacidades de sombreado en la figura 4. Un segmento inicial de la sílaba A (marcado con sombreado más claro en la figura) se utiliza para detectar el propósito de producir un canto en el gesto de presión del ave enmudecida. La parte que se produce luego del disparo de la síntesis aparece en un tono más oscuro del mismo color. Estas grabaciones fueron utilizadas para ajustar los parámetros del modelo para producir vocalizaciones sintéticas que presentan una coincidencia en la frecuencia fundamental y el contenido espectral.

El animal es luego enmudecido y colocado en la caja acústica y conectado para controlar el órgano vocal electrónico a través de su gesto de presión.

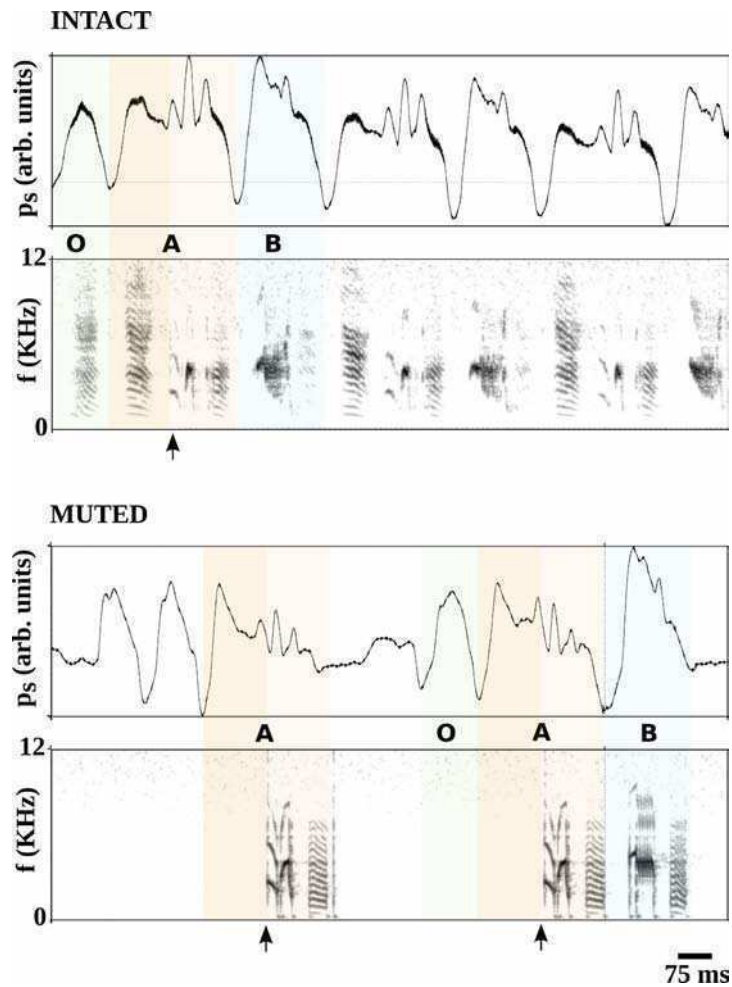


FIGURA 4. Canto intacto y canto sintético, controlado por el sujeto. Gesto de presión y sonograma, con diferentes colores y opacidades de sombreado indicando las distintas sílabas, y una flecha marcando el segmento de la primer sílaba utilizada para la detección (paneles superiores). Cuando el pájaro enmudecido conduce la siringe, se observa en el sonograma que canto sintético es producido luego de que la primera sílaba es detectada y hasta que se reconoce la interrupción del motivo (paneles inferiores).

En el panel inferior de la figura 4 se muestra el gesto de presión del ave enmudecida y el sonograma del registro de un micrófono colocado cerca (a unos 20 cm) del pájaro y el parlante. Se puede observar en el sonograma que no se produce sonido en los intentos del ave de pronunciar una nota introductoria. Cuando el gesto de presión corresponde a la primera sílaba en el motivo (sílabas A), la instrucción es reconocida y el canto correspondiente es sintetizado y reproducido a través del parlante. En la primera frase, el pájaro solo evoca el gesto correspondiente a la primera sílaba de un motivo, y luego se detiene. En este caso, el algoritmo detecta la interrupción y se

---

apaga la integración. En la segunda frase, el pájaro sigue con la segunda sílaba y conduce la siringe electrónica hasta el final del motivo.

Durante diez horas de experimento, un pájaro enmudecido intentó alrededor de 450 llamadas, de las cuales menos de 5% generaron falsos disparos de canto sintético. En la mayoría de los casos la activación de eventos falsos fue reconocida por el algoritmo y silenciada después de menos de 20 ms. La tasa de éxito en la detección del inicio de un motivo fue de alrededor de 65% en 150 intentos de canto realizados por el ave.

#### 4. Perspectivas para BMI

Mostramos aquí que comportamiento vocal realista se sintetiza *on-line* por nuestro dispositivo, al ser controlado por el gesto motor espontáneo de un pájaro enmudecido, una señal fisiológica (la presión del saco aéreo) que está degradada con respecto a la registrada en pájaro intacto. La plataforma de cómputo es un procesador portátil de bajo costo, y la tasa inicial de éxito es alta. Este es un ejemplo alentador de la plausibilidad de una especie de interfaz entre el generador central de patrones motores y comportamiento sintético bio-mimético. La tecnología DSP está siendo implementada en una variedad de problemas inspirados en la biología, y junto con la tecnología Field Programmable Gate Array (FPGA) es candidata a convertirse en una solución estándar para una gran variedad de aplicaciones bio-miméticas [SCWM06, DCM+06].

Las BCI y BMI suelen leer datos fisiológicos y tratar de descifrar instrucciones motoras para conducir dispositivos periféricos con el fin de producir una retroalimentación sensorial sintética [SCWM06, DCM+06, CMM+99, CLC+03]. Porque tenemos un modelo físico del efector periférico, los orígenes de la complejidad del comportamiento están vinculados a caminos suaves en un espacio de parámetros de baja dimensión. Tras la identificación de los parámetros pertinentes y su vínculo fisiológico con el generador de patrones (actividad vS y presión del saco aéreo sub-siríngeo), una mayor simplificación del sistema se llevó a cabo, con bases dinámicas, mediante la eliminación de términos no lineales irrelevantes (realizando una reducción del modelo a su forma normal). Esto dio lugar a la posibilidad de implementar nuestro dispositivo bio-protésico en una plataforma electrónica programable. Dado que las capacidades de cómputo de la plataforma se ven incrementadas considerablemente por el bajo costo de nuestra implementación, avances tecnológicos en este frente tendrán un impacto amplificado en la complejidad del sistema periférico biomecánico que puede ser emulado.

Parte 4

Desenlace y final



## Conclusiones

### 1. Objetivos cumplidos

#### 1.1. Interacciones en la periferia.

*El fenómeno.* Estudiamos la dinámica de un sistema sencillo formando por una fuente de sonido que interactúa acústicamente con un tracto vocal. Tomamos como motivación la presencia de fenómenos acústicos complejos no correlacionados con instrucciones motoras complejas, como los saltos de frecuencia que ocurren unilateralmente en el canto del ruiseñor [ZRS08]. Ha sido reportado además que algunos oscinos ajustan la geometría del tracto vocal en coordinación con la variación de la frecuencia fundamental de las vocalizaciones [RSFB06, FRS06].

Hemos mostrado que, para longitudes del tracto tales que sus resonancias son cercanas a la frecuencia fundamental del sonido producido en la fuente, existe una multi-estabilidad de soluciones periódicas. Este escenario dinámico implica que son posibles cambios rápidos en la acústica del sistema, incluso con ligeras variaciones de los parámetros (asociados a gestos motores).

El sistema dinámico con el que se modela el órgano vocal presenta poca complejidad. En contraste con otras descripciones de la dinámica de las membranas oscilantes en la fuente de sonido de humanos o de aves, como el de dos masas, el modelo de aleteo extendido presenta un único modo de oscilación espacial. La simplicidad del modelo que propusimos para estudiar el fenómeno es tal que sólo presenta una bifurcación de Hopf como mecanismo de inicio de oscilaciones labiales.

Al introducir la interacción con el tracto se genera comportamiento complejo. En nuestro modelo, el tracto vocal no actúa simplemente como un filtro que resalta ciertas frecuencias y atenúa otras [ML05], sino que interactúa con el aparato fonador mediante la inducción de una realimentación retardada en la presión supra-siríngea  $p_i$ . Observaciones recientes indican que mediante la variación de la geometría del tracto vocal, es posible ajustar la realimentación en la presión inter-labial [HFH06]. Dentro del escenario complejo de bifurcaciones al que lleva la introducción de la interacción fuente-tracto, esta realimentación retardada aparece como un candidato para la producción de saltos en la frecuencia de las vocalizaciones.

*La estrategia.* La minimalidad del modelo para la fuente de sonido persigue dos objetivos. Por un lado, los efectos no lineales que se introducen en la dinámica y que se reflejan en las propiedades espectrales y temporales de las vocalizaciones, pueden ser identificados con la interacción acústica con el tracto.

---

Por otro lado, el acople con el tracto, aún modelado en una primera aproximación como un tubo de sección uniforme, plantea un sistema de ecuaciones diferenciales con retardo. Este tipo de sistemas es difícil de tratar. En esta parte, el hecho de que las soluciones de oscilador libre de la fuente sean las de un oscilador en dos dimensiones facilita el tratamiento analítico mediante las aproximaciones adecuadas.

*Los hallazgos.* Al llevar al sistema tracto-fuente al límite de disipación infinita, la dinámica altamente no lineal de la fuente se reduce, para las soluciones oscilatorias, a trayectorias sobre un ciclo límite estable. Esas trayectorias pueden ser descritas en términos de una variable de fase unidimensional. Al considerar el efecto de la fuente, se introduce una perturbación a la dinámica de la fase. De este modo fuimos capaces de determinar bajo qué condiciones es posible la coexistencia de soluciones periódicas. Esta coexistencia, proponemos, es un mecanismo plausible para la producción de vocalizaciones con un salto en la frecuencia, requiriendo apenas al sistema de control (el sistema nervioso central) un gesto motor suave: un ajuste de la longitud del tracto durante la fonación.

La coordinación de las actividades de la fuente y el tracto no siempre conduce a características acústicas observables en la frecuencia (como saltos o subarmonicidad) [FRS06, Now87]. Debido a esto, tomamos también un camino analítico en busca de trazas del acople fuente-filtro en la amplitud de los sonidos. En el límite débilmente no lineal, donde las soluciones del sistema son pequeñas desviaciones respecto de oscilaciones armónicas, estudiar las bifurcaciones de los ciclos límite resulta equivalente a estudiar las de los puntos fijos de las ecuaciones que rigen la dinámica de las desviaciones. Así, pudimos determinar que la introducción de la interacción con el filtro introduce en el sistema la posibilidad de que ocurra una bifurcación de tipo *cusp* en los ciclos límite. Como consecuencia, las resonancias del sistema fuente-filtro se intensifican y se desplazan, se crean regiones de coexistencia y se cambian los umbrales de fonación.

## 1.2. La siringe electrónica.

*La aplicabilidad del modelo.* Nos apoyamos en resultados previos que siguen esta línea de pensamiento, la que resalta que para la periferia es posible producir comportamiento complejo requiriendo poca complejidad por parte de las instrucciones con las que se la conduce. Tomamos los avances en el modelado del órgano fonador del diamante mandarín [SAGM08, AM08], y los aplicamos a la construcción de un dispositivo bio-protésico.

Identificamos los parámetros que determinan las propiedades espectrales de soluciones periódicas en función de la distancia a determinadas bifurcaciones. Cerca de una bifurcación de Hopf, las oscilaciones son tonales. Cerca de una bifurcación SNIC, las oscilaciones mantienen una traza de la existencia de puntos estacionarios, y son espectralmente ricas. Los parámetros que determinan esas propiedades en el modelo están relacionados con la actividad del músculo *vS* y con la presión sub-siríngea.

Al proponer un conjunto de series temporales adecuadas en estos parámetros, pudimos generar canto sintético realista. En una primera aproximación, bastó con un modelo de aleteo extendido para la fuente y un modelo muy

---

sencillo de tracto no interactuante, que incluye apenas un tubo emulando la tráquea.

*La factibilidad de la implementación.* El modelo, aún de baja dimensión, pudo ser simplificado. Lo llevamos a la forma normal alrededor de la bifurcación de Takens-Bogdanov. Es decir, realizamos una eliminación de términos no lineales que son irrelevantes en la determinación de los mecanismos dinámicos que dan origen a las oscilaciones cualitativamente distintas en las membranas de la siringe.

En esta forma el modelo es computacionalmente más simple y sugiere aún que el control de los parámetros determinantes de la dinámica de las oscilaciones puede efectuarse mediante registros fisiológicos. Obtuvimos esos registros. Con un transductor piezoeléctrico registramos la presión en el saco aéreo y mediante el implante de electrodos en la masa del músculo vS registramos su actividad EMG, en simultáneo con el canto.

Encontramos transformaciones que llevan de los registros fisiológicos a parámetros del modelo, de modo que el canto sintético que se obtiene al integrar numéricamente es comprable con las grabaciones. Pudimos llevar a cabo la integración en tiempo real a partir de gestos motores fisiológicos reales, programando la transformación y las ecuaciones del modelo en una plataforma portátil de procesamiento de señales digitales.

*Un prototipo exitoso.* Habiendo construido confianza en la factibilidad de la integración en tiempo real del modelo conducido por registros fisiológicos, nos aventuramos en la construcción de un dispositivo que emula comportamiento vocal complejo cuando es controlado por un sujeto mediante su gesto de presión. Las vocalizaciones realistas sintéticas se producen *on-line*, mediante el cómputo del modelo del órgano vocal en un procesador de señales digitales.

Para incrementar el realismo de la síntesis, agregamos más detalle al modelo del órgano vocal del diamante mandarín, incluyendo una descripción física de la glotis, la OEC y el pico. En la OEC, una gran masa de aire actúa de forma inercial ante las perturbaciones acústicas generadas en la fuente y propagadas a lo largo del tracto. En nuestra descripción, la OEC funciona como un resonador de *Helmholtz*, cuyo cuello actúa del mismo modo que la glotis. Esta parte del órgano fonador pudo ser descrita, como es usual en acústica, a través de un circuito eléctrico equivalente, y de este modo mediante un sistema dinámico (el que da cuenta de la evolución temporal del flujo de aire y la presión acústica, análogos a la corriente y el potencial eléctrico). Lo que se consigue al incrementar el nivel de detalle en la descripción del tracto es un trazo más fino en la composición espectral de los sonidos sintéticos. Un segundo DSP fue dedicado al cómputo del sonido que emerge del tracto, a partir de las perturbaciones acústicas generadas en la fuente.

El dispositivo, conducido por el gesto de presión de un diamante mandarín enmudecido comportándose libremente, tuvo que contar con un algoritmo de detección y selección. A partir del gesto motor degradado, se midió en tiempo real la correlación con un segmento del gesto de presión registrado en el pájaro intacto. Esta parte, que es semejante a la etapa de extracción de características en los registros neurofisiológicos con los que se



---

controlan las BCIs, se resolvió con computación muy sencilla dado que el gesto motor está adecuadamente identificado mediante el modelado de la interfaz, y es una forma estereotipada en un espacio de baja dimensión.

El prototipo es exitoso: el motivo de canto se sintetiza y se reproduce en tiempo real bajo el control fisiológico de un pájaro previamente enmudecido comportándose libremente, reemplazando la realimentación auditiva propia. Este éxito construye confianza en la plausibilidad de un tipo de interfaces cerebro-comportamiento en las que el modelado de la periferia es clave para evitar la estadística en registros extensivos de actividad cerebral.

## 2. Aprendimos algo, hoy

### 2.1. Complejidad en la periferia vs. complejidad en el cerebro.

Los resultados de esta tesis aportan a la idea de que el comportamiento complejo tiene al menos parte de su origen en la periferia biomecánica, tomando como modelo animal a los oscinos y su canto [ZM04, ML05]. La producción de vocalizaciones de contenido espectral diverso se viene mostrando relacionada más con las posibilidades del órgano fonador que con un control motor sofisticado; bien porque la física de las membranas oscilantes en la siringe no requiere más que instrucciones suaves, o porque al tener en cuenta las posibles interacciones entre sistemas complejos la periferia es capaz de producir acústica rica aún bajo control pasivo.

*Interacción entre sub-sistemas periféricos.* La introducción de interacciones fuente-filtro agrega complejidad, al introducir una dinámica de infinitas dimensiones en las ecuaciones con las que se modela el órgano fonador. Esta complejidad se manifiesta en efectos acústicos. Lo que nos motiva a buscar el origen de complejidad acústica en la interacción fuente-tracto son observaciones experimentales: las aves acompañan el conjunto de gestos motores (sencillos) con los que controlan la siringe con gestos motores (sencillos) coordinados para controlar la geometría del tracto en sintonía con determinadas propiedades de los sonidos producidos en la fuente.

El método para construir confianza sobre nuestras hipótesis acerca del origen de la complejidad en el comportamiento fue: proponer un sistema cuya dinámica no lineal es conocida y sus bifurcaciones sencillas; introducir en la descripción la interacción con otro subsistema, y estudiar la dinámica del sistema completo. En la dinámica resultante, es fácil determinar de qué efectos es responsable la introducción de la interacción (en este caso, la de la fuente con el tracto). Reduciendo el sistema mediante aproximaciones, se determina cuáles son los mecanismos dinámicos que subyacen a la complejidad que deriva en la aparición de propiedades acústicas no triviales. De este modo fuimos capaces de proponer de qué modo podrían generarse propiedades acústicas no triviales en la fonación, observadas tanto en la frecuencia (saltos) como en la composición espectral del canto (enfaticación y desplazamiento de las resonancias). Llegamos a proponer experimentos en los que se puede cuantificar en qué grado la fuente y el tracto están acoplados acústicamente.

*El canto del diamante mandarín.* El diamante mandarín posee un canto complejo, que cubre un rango amplio de frecuencias y contenidos espectrales. Una versión del modelo de aleteo extendido para la siringe consigue

---

proponer mecanismos intrínsecos a la biomecánica que dan cuenta de esta diversidad. Las propiedades acústicas de los sonidos que las aves emiten pueden relacionarse con características de las soluciones posibles de un sistema dinámico determinado [SAGM08, AM08].

Más allá de esa identificación, al vincular parámetros del modelo con gestos motores fisiológicos registrados en un ave en nuestro laboratorio, pudimos generar canto sintético realista [SAGM10, PAA<sup>+</sup>11, APGM12]. Esos gestos motores son simples: representan caminos suaves en un espacio de parámetros del modelo de apenas dos dimensiones.

**2.2. El modelo como camino.** Trabajamos con un modelo para la fuente que presenta un único modo espacial y apenas dos parámetros de control activo para determinar propiedades acústicas de sonidos emergentes, mayormente la frecuencia fundamental y el contenido espectral. Para estudiar la complejidad de un sistema fuente-filtro, propusimos un modelo muy sencillo para la fuente y también para el tracto vocal: tuvimos en cuenta la tráquea y la aproximamos por un tubo, si bien el tracto vocal superior incluye a la glotis, la cavidad oro-faríngea-esofágica, y el pico [FRS06, RSFB06].

Sin embargo, estos niveles de descripción mostraron ser adecuados para comprender los mecanismos subyacentes al comportamiento complejo y para identificar los parámetros fisiológicamente relevantes. El trabajo de buscar esta jerarquía de los fenómenos que ocurren en un sistema tan complejo como un órgano fonador, deja como saldo un modelo que avanza iterativamente en la comprensión de la interfaz entre el sistema nervioso central y las ricas vocalizaciones de los oscinos [GS95, GS96, LGM02, SAGM08, AM08, AM09, PAA<sup>+</sup>11]. La comprensión del funcionamiento de esta interfaz en términos de la Dinámica No Lineal de los sistemas bio-físicos interactuantes, tiene además la capacidad predictiva que se ha mostrado, por ejemplo, en esta tesis.

La dinámica del aparato fonador sugiere además mecanismos de control, orientando las exploraciones sobre cómo es la interacción de la periferia con el sistema nervioso central, sobre cuál es la naturaleza y la forma de las instrucciones con las que el generador de patrones puede conducir la periferia para producir el canto. La comprensión de la no linealidad del efector periférico mediante un modelo implica además la comprensión de la relación no trivial (no lineal) entre los gestos motores y el comportamiento.

**2.3. El cuerpo manda.** Si esa relación gesto motor/comportamiento está mediada por un efector no lineal, ciertas restricciones en el comportamiento pueden estar vinculadas a propiedades del efector, y no necesariamente a una selectividad en el sistema donde se originan los patrones motores.

En el canto del diamante mandarín, el contenido espectral de un sonido no es independiente de la frecuencia fundamental. En el capítulo 6 produjimos sonidos integrando el modelo de la siringe del diamante mandarín para una grilla de parámetros identificados con el control motor. Estos sonidos reprodujeron la dependencia contenido espectral/frecuencia fundamental, mostrando que esa restricción viene impuesta por la biofísica de la fuente.



## Perspectivas

### 1. Todavía más complejidad en la periferia

**1.1. Interacción acústica entre las dos fuentes.** Así como el acople entre la fuente y el filtro es causa de complejidad y fenómenos no lineales adicionales a los que ocurren en la fuente, la interacción entre las dos fuentes de sonido tiene el potencial de introducir bifurcaciones que originen efectos en la frecuencia de las vocalizaciones.

Si se tienen en cuenta dos fuentes simples, modeladas del mismo modo que la fuente introducida en el capítulo 3, ellas quedan acopladas entre sí al considerar la presión a la entrada del tracto, cuando se contabiliza la presión inter-valvular a partir del flujo de aire. La razón de esto es que la presión en la entrada del tracto ahora recibe contribuciones de las fluctuaciones en ambas fuentes:

$$p_i(t) = \alpha\sqrt{p_s}(x_1 + x_2) + \alpha\sqrt{p_s} \sum_{n=1}^{\infty} (-r)^n [x_1(t - n\tau) + x_2(t - n\tau)], \quad (40)$$

donde  $x_{1,2}$  son las desviaciones de los puntos medios de los labios de cada una de las fuentes, y el resto de los coeficientes son los mismos que en la ecuación (5).

Aún sin considerar las reflexiones en el extremo abierto del tracto, el sistema resultante es uno de dos osciladores acoplados. Despreciando la contribución de las reflexiones (con  $r = 0$ ), el sistema de ecuaciones para las membranas oscilantes de ambas fuentes (análogas a la ecuación (21)) en la siringe es

$$\begin{cases} \dot{x}_1 + \omega_{01}^2 x_1 &= \mu [x_1 - \tilde{c}x_1^2 x_1 + (1 - \dot{x}_1)\tilde{\alpha}(x_1 + x_2)], \\ \dot{x}_2 + \omega_{02}^2 x_2 &= \mu [x_2 - \tilde{c}x_2^2 x_2 + (1 - \dot{x}_2)\tilde{\alpha}(x_1 + x_2)]. \end{cases} \quad (41)$$

Este sistema de dos osciladores acoplados puede ser estudiado en el límite débilmente no lineal, para encontrar analíticamente bifurcaciones en las soluciones que puedan vincularse con las propiedades acústicas de los sonidos producidos [DK91]. En particular, las diferencias de fase posibles entre los dos osciladores determinan efectos asociados a cambios en la frecuencia (duplicación, por ejemplo). Cambios en la estabilidad de soluciones correspondientes a distintas diferencias de fase conllevan fenómenos introducidos por el acople, y no necesariamente dependientes del control muscular independiente de las fuentes [ZS04, LSZM08].

### 2. Siringe electrónica

Los avances en la comprensión y el modelado matemático del funcionamiento del órgano fonador conducirán a mejorar la síntesis alcanzada por

---

los próximos prototipos de la siringa electrónica. Así como la descripción más detallada del tracto vocal devino en una síntesis más realista sin requerir un cambio de plataforma computacional ni registros adicionales para el control, los descubrimientos iterativos de efectos acústicos serán aplicables dentro del modo actual de operar la prótesis. Entre estos, los favoritos son los relativos a las interacciones acústicas entre los subsistemas de la periferia, que son los candidatos a producir cambios cualitativos sin necesidad de control sofisticado.

Al estar basada en el modelado, la siringa electrónica que presentamos también abre camino a una serie de experimentos que exploren la función de la realimentación sensorial auditiva en la adquisición del canto. Al mismo tiempo, lo que construimos es un prototipo de una BMI que produce canto sintético bajo el control de un sujeto, a partir de un gesto motor de baja dimensión y con una plataforma de cómputo en tiempo real portátil estándar.

**2.1. Experimentos con realimentación sensorial.** Uno de los resultados del programa del que esta tesis forma parte, es que hay restricciones en las propiedades acústicas del canto del diamante mandarín (la relación frecuencia fundamental/contenido espectral) que vienen determinadas en paquete por la periferia. Esto motiva el diseño de experimentos para determinar en qué grado esta restricción en el espacio de propiedades acústicas posibles en la periferia restringe la libertad del sistema nervioso central en la exploración de patrones motores durante el aprendizaje/reaprendizaje.

Contar con esta observación, con un modelo predictivo y con la siringa electrónica, nos habilita para una exploración del rol de la realimentación sensorial para la construcción de confianza en una de dos hipótesis que compiten: o bien la adquisición y corrección de los gestos motores que se corresponden con las vocalizaciones del adulto se da mediante una búsqueda aleatoria en un espacio amplio de instrucciones, o esta búsqueda está restringida por las características de la periferia.

En los experimentos de realimentación sensorial alterada que proyectamos, la forma de producir las alteraciones debe estar correlacionada con alteraciones en las instrucciones motoras. Un protocolo experimental para atacar la pregunta planteada en el párrafo anterior estará basado en el estudio de los mecanismos de reaprendizaje de las instrucciones motoras cuando un ave es enmudecida y su canto es reemplazado por canto sintético generado en tiempo real y controlado por sus propias instrucciones fisiológicas. Al alterar las instrucciones artificialmente y en sintonía con alteraciones en los parámetros que identificamos relacionados con los gestos motores del ave, los mecanismos de reaprendizaje podrán ser estudiados. En particular, si la respuesta del ave es compensatoria, construiremos confianza sobre la hipótesis de que las restricciones de la periferia al espacio de las propiedades acústicas impone condiciones a las exploraciones en el espacio de las instrucciones motoras.

---

### 3. BMI

La evolución de las BMI llegó a la construcción de dispositivos que permiten que un primate controle un brazo mecánico para alcanzar y agarrar objetos [CLC<sup>+</sup>03, LN06, SCWM06]. A las interfaces de este tipo se las llama BMI de circuito cerrado (*closed loop BMI*) porque hay un efector artificial que proporciona al sujeto una realimentación sensorial (ver el brazo moverse, alcanzar con él un objeto y tomarlo, etc). Lo usual en este tipo de implementaciones es registrar algún tipo de actividad cerebral en simultáneo con la realización de la tarea. La mayoría registran electrodos extra-celulares que miden la actividad de un número que va de decenas a centenas de unidades en corteza motora [CLC<sup>+</sup>03]. Hay también técnicas no invasivas, que se basan en otras señales como EEG y fMRI [CWR<sup>+</sup>08]. Sobre el conjunto de registros se hace estadística y se aplican algoritmos de aprendizaje automático, para clasificar ciertas características que vinculen la actividad en el cerebro con una intención motora determinada [CMM<sup>+</sup>99, SCWM06].

Mientras los dispositivos intentan “aprender” esas relaciones causales, el sujeto aprende a controlar el dispositivo para obtener una determinada realimentación. Este aprendizaje se evidencia en el cambio de las características presentes en los registros a lo largo de las repeticiones de la tarea [CLC<sup>+</sup>03, CWR<sup>+</sup>08]. Mediante estas interfaces se pretende contribuir a la comprensión de los orígenes de las instrucciones motoras y de los vínculos entre la plasticidad sináptica y el aprendizaje motor [LN06].

**3.1. Una estrategia intermedia.** En trabajos recientes llevados a cabo en el instituto de rehabilitación de Chicago (*Rehabilitation Institute of Chicago*) se utilizó actividad *EMG* para controlar extremidades protésicas implantadas en humanos [KLL<sup>+</sup>09]. En esas implementaciones, nervios residuales de extremidades amputadas son transferidos quirúrgicamente a sitios alternativos de tejido muscular. Luego la actividad de estos músculos es registrada y utilizada para el control de las prótesis.

Se utilizan 12 electrodos bipolares para registrar la actividad *EMG* en los músculos a los que se re-dirigieron los nervios. Sobre estos registros se aplican técnicas estadísticas semejantes a las empleadas por las BMI mencionados más arriba, para extraer características y clasificar patrones de actividad correspondientes al control de hasta 10 grados de libertad de movimiento de un efector protésico mecánico.

Aparte de la ventaja de que la actividad *EMG* es más sencilla de registrar, el conjunto de registros es de dimensión más baja (al menos un orden de magnitud menor). Por esa razón, la clasificación requiere relativamente mucha menos capacidad de cómputo que en otras implementaciones de BMI.

**3.2. Qué resuelve el modelo de la periferia.** La etapa crítica de muchas implementaciones es la de extracción de características y clasificación (*feature extraction and classification*), que se debe realizar en tiempo real sobre un conjunto grande de registros. Nuestra implementación facilita la selección de características, porque se hace en un espacio de parámetros fisiológicos de baja dimensión, pero además porque no se propone un filtro entre la actividad neuronal y el comportamiento, sino que hay un modelo

---

que predice, a partir de los valores de los parámetros, una respuesta no lineal del efector periférico.

El aporte de las BMI actuales a la comprensión del origen de las instrucciones motoras se apoya en que las características que se extraen tienen potencialmente un sentido fisiológico. La separación de fuentes a ciegas (*Blind Sources Separation*), por ejemplo, asume que en cada uno de los registros se mide la suma lineal de las contribuciones de un número finito de fuentes [CZPA09]. En este paradigma, la clasificación de características implica invertir esa combinación lineal y separar las contribuciones de cada una de las fuentes que corresponden a una determinada tarea motora [LCiA04, PS00]. Al hacer esta descomposición, se identifica un número de fuentes no estacionarias que contribuyen a un gesto motor. Estas fuentes, en principio abstractas, pueden aportar a la tarea de identificar los orígenes de las instrucciones motoras a partir de registros de, por ejemplo, *EEG* [LKCC07].

En nuestra implementación, es el mismo modelo de la periferia el que sugiere los parámetros determinantes del comportamiento. Encontrar el sentido fisiológico de esos parámetros, si bien no es una tarea trivial, es una que está orientada por la dinámica no lineal de los sistemas biofísicos que actúan. Esto es cierto tanto para su naturaleza como para su comportamiento. Al momento de construir la siringe electrónica para el diamante mandarín, hicimos pie en que los parámetros determinantes de propiedades acústicas como la frecuencia fundamental y el contenido espectral provienen de la presión en los sacos aéreos y de la actividad del músculo vS. Estudiando en más detalle la dinámica del modelo, encontramos que con caminos suaves en este espacio de parámetros se produce canto sintético comparable con el real. En [PAA<sup>+</sup>11] mostramos en detalle cómo estos caminos son apenas transformaciones lineales del conjunto de los dos registros fisiológicos.

### 3.3. Resultados hermanos.

*3.3.1. Realismo según el pájaro.* El modelado de la biofísica de la periferia implicó una jerarquización de los fenómenos, la elección de una escala de descripción. El resultado es que podemos sintetizar canto con dos registros fisiológicos, y que ese canto es comparable con el real. En esta tesis, a la comparación entre la predicción del modelo y el comportamiento observado la hacemos mayormente en términos de la distancia (euclídea) en el espacio de dos dimensiones que comprenden dos características acústicas que cuantificamos: la frecuencia fundamental y el contenido espectral. Hasta ahora queda afuera la percepción del sujeto acerca del realismo de nuestra síntesis.

Con la siringe electrónica construida, apostamos a la realización de experimentos de realimentación sensorial alterada en los que responderemos, con experimentos comportamentales, preguntas acerca del rol de la dinámica de la periferia en los mecanismos de aprendizaje de las instrucciones motoras. Una de las primeras preguntas es justamente si la realimentación sintética suscita la misma respuesta que la realimentación propia.

Hay un resultado de *Amador, Sanz Perl, Margoliash y Mindlin* que construye confianza en el realismo de nuestro canto sintético. Algunas neuronas en el sistema vocal responden selectivamente al presentarle al un ave estímulos auditivos: responden ante la presentación de grabaciones del canto propio

---

pero no, por ejemplo, a cantos con-específicos, o al canto propio invertido [Mar86]. En experimentos recientes, el canto sintetizado por nuestro modelo obtuvo el mismo tipo de respuesta que el canto propio [APMM].

3.3.2. *Baja dimensión, también en el cerebro.* Los patrones de presión que hacen de control en el sistema no lineal implicado por la periferia son suaves y estereotipados. Esta observación conduce a la hipótesis de que los mecanismos de control en el sistema nervioso central son simples: el modelado de la interfaz entre la actividad neuronal y el comportamiento asiste en la proposición de modelos de funcionamiento del sistema de canto [GCM01].

Asumiendo que la actividad respiratoria que se registra es función de un observable que se proyecta desde el sistema vocal, en la tesis de doctorado de *Leandro Alonso*, también del Laboratorio de Sistemas Dinámicos ([AAAM08, AAGM09, AAM10]), se explora la hipótesis de que los patrones motores emergentes del sistema nervioso central se pueden entender en términos de la actividad de un sistema de baja dimensión.

En esos trabajos, se propone que las actividades de sub-poblaciones de neuronas excitatorias e inhibitorias interactúan a la manera de un oscilador de *Wilson-Cowan* [WC72], en un núcleo cuya actividad promedio es la responsable de controlar los músculos respiratorios que originan los patrones de presión. Este núcleo, modelado en última instancia como un oscilador no lineal de dos dimensiones, es forzado por la actividad de otro, ubicado en una mayor jerarquía en el sistema de canto. Tanto en canarios (*Serinus canaria*) como en el diamante mandarín, los gestos de presión registrados pueden ser reproducidos por este sistema de osciladores neuronales.

Al emplear la siringa electrónica basada en un modelo de la periferia, las conclusiones relacionadas con el origen de las instrucciones motoras y su plasticidad podrán vincularse a la actividad neuronal mediante modelos de baja dimensión. Así como se puede detectar el gesto motor para reproducir el canto, se puede emplear el mismo gesto motor para ajustar los parámetros de una arquitectura neuronal sencilla, como la propuesta en los trabajos mencionados más arriba.

Respecto de los métodos estadísticos utilizados por las BMI actuales para inferir los orígenes del control motor en fuentes abstractas, los modelos propuestos en [AAGM09, AAM10] aportan, sobre todo, capacidad predictiva. Una vez ajustados los parámetros que dan cuenta de la arquitectura de conexiones, a partir de la actividad registrada en la periferia se pueden reconstruir otros observables del sistema.

#### 4. Jugar en toda la cancha

Nuestra implementación de la siringa electrónica se programó en un chip *TMS320C67x*, introducido en 2002 por *Texas Instruments* [Keh05]. Desde entonces, los *chips* programables empleados para el procesamiento de señales han virado a combinaciones de procesadores dedicados con lógica programable, como los *FPGA* (*Field Programmable Gate Array*) [MB07]. Así como los DSP optimizan la computación de señales porque tienen bloques de *hardware* dedicados, los FPGA permiten sintetizar bloques de lógica a la medida de los algoritmos. De este modo, distintas etapas de la integración de los modelos pueden implementarse como bloques de circuito que funcionan en



---

paralelo. Una ventaja adicional es que este modo de llevar a cabo la computación de señales en tiempo real se está convirtiendo en un estándar de las comunicaciones (en teléfonos celulares, en computadoras portátiles, en GPS), por lo que se especula que la capacidad de esta plataforma crecerá de modo sostenido.

El modelo de aleteo extendido para la siringe de los oscinos presentado hace diez años en [GCM01] como una apuesta a capturar la física de la periferia relevante para producir un comportamiento, hoy sintetiza canto realista a partir de registros fisiológicos, y es el corazón de una siringe electrónica controlada por un sujeto enmudecido. Este dispositivo va a permitir explorar más profundamente los mecanismos que originan el canto, que serán a su vez aplicables a la mejora iterativa del prototipo.

Dentro de un programa más amplio, esa exploración puede abarcar la comprensión de los mecanismos de generación de las instrucciones motoras en el cerebro, también mediante modelos matemáticos de baja dimensión. Si se tiene éxito en esto y se consigue vincular el comportamiento con los (pocos) parámetros de estos modelos, la etapa de clasificación de características en los registros fisiológicos puede mejorarse cualitativamente.

A medida que se incrementa la capacidad de hacer computación en tiempo real, la comprensión de los sistemas periféricos mediante el modelado de la biofísica se presenta como una propuesta alternativa a los métodos estadísticos en espacios de parámetros de alta dimensión.

## Circuito para rectificación y suavizado de EMG

### 1. Rectificación y amplificación

Medimos la actividad del músculo mediante la inserción de un par de electrodos en la masa muscular, que registran la actividad eléctrica de un número de fibras. Esta actividad eléctrica *EMG* es luego rectificada y suavizada. Hacemos esto siguiendo un modelo propuesto en [RBF<sup>+</sup>97].

Si  $vS_{raw}(t)$  a la señal *EMG* de entrada, la actividad suavizada y rectificada  $vS$  se obtiene de

$$\dot{vS} = -\lambda vS + |vS_{raw}(t)|^2. \quad (42)$$

### 2. Implementación electrónica analógica

Construimos un circuito que presenta una dinámica equivalente La señal

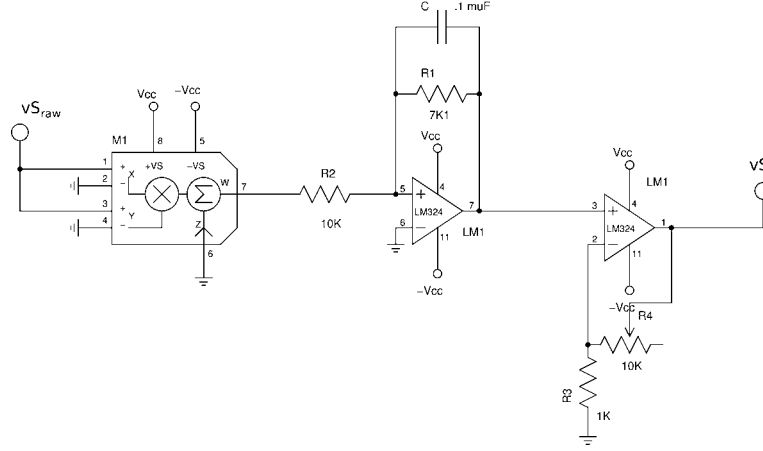


FIGURA 1. Circuito de rectificación, promediado y amplificación.

ingresa y es elevada al cuadrado en un AD633. Como las patas + y - del primer amplificador operacional están a la misma tensión, puede calcularse fácilmente que (si llamamos  $V_i$  a la entrada y  $V_o$  a la salida de esta etapa del circuito)

$$\dot{V}_o = -\frac{V_o}{R_1 C} - \frac{V_i}{R_2 C}. \quad (43)$$

Teniendo en cuenta que  $V_i = \frac{|vS_{raw}|^2}{10}$ , la ecuación del circuito es equivalente a (42).



## Amplificador de la señal de presión

### 1. Diseño del Amplificador

**1.1. Alimentación del Fujikura.** Se indentificó como una fuente de variabilidad en la medición posibles inestabilidades en la corriente de alimentación del sensor. Según la hoja de datos del *Fujikura FPM-02PG* debe circular  $1,5\text{ mA}$  de corriente, por lo que se diseñó, como muestra la figura 1, un circiuto capaz de fijar de forma precisa ese valor de corriente. Para ello se diseñaron dos etapas: en una se utilizó un fijador de voltajes de referenciaa regulable, el *LM385* como fuente de tensión lo suficientemete precisa capaz de entregar ese valor de corriente. Así se fijó un  $V_{ref} = 1,24\text{ mV}$  para que sirva de fuente de tensión para la siguiente parte del circuito. Ésta se encarga de brindarle la corriente necesaria al sensor y consta de un *OP290* realimentado negativamente y figura en la hoja de datos del *Fujikura* como un esquema de posible.

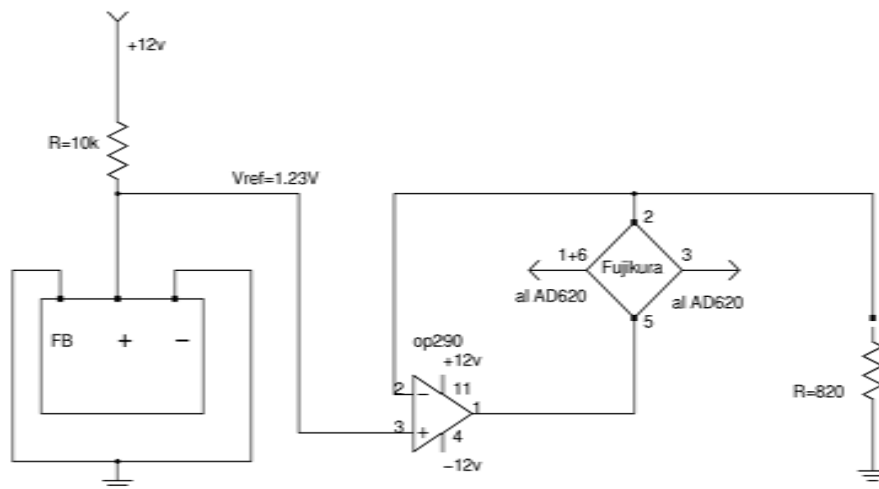


FIGURA 1. Esquema de la etapa de Alimentación del Fujikua

**1.2. Amplificación y Filtrado.** La señal que mide el *Fujikura* es diferencial, en el sentido que se obtiene a partir de realizar la diferencia entre dos de las terminales del sensor. Por ello se determinó emplear un amplificador diferencial *AD620* para en un sólo integrado tener la diferencia amplificada. Esta amplificación se mantiene fija, ya que es posible que la

señal contenga un nivel de continua por lo que una amplificación variable implicaría ir modificando este valor dificultando las siguientes etapas. Luego la diferencia amplificada, que representa ya la señal de presión buscada, se filtra con un pasa-altos tipo Bessel de orden 2 con una frecuencia de corte muy baja ( $f_c = 0,01 \text{ Hz}$ ) con el fin de eliminar el nivel de continua, los valores de los distintos componentes se muestran en el esquema de la figura 2

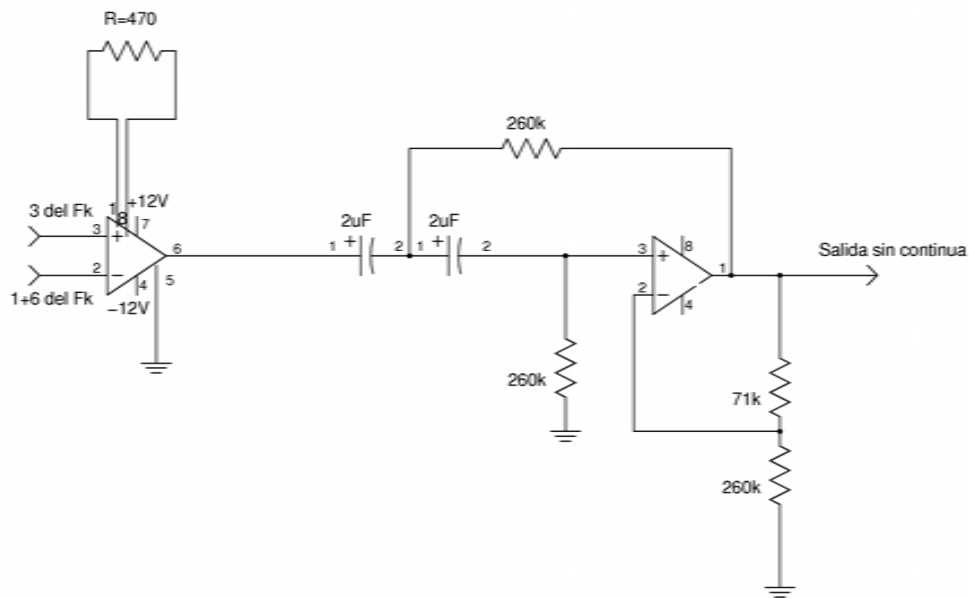


FIGURA 2. Esquema de la etapa de amplificación y filtrado de la señal. A la izquierda del circuito se muestra la amplificación diferencial. A la derecha se muestra el esquema del filtro pasa-altos.

**1.3. Control de Nivel de Continua.** La última etapa del amplificador se muestra en la figura 3, consta de un amplificador inversor, un seguidor con amplificador y de un sumador que hace la suma pesada de la salida de ambos amplificadores. El seguidor toma un voltaje de referencia  $V_{ref} = 0,05 \text{ V}$  de un circuito como el de la figure 1 más un divisor de tensión y amplifica esa señal con una relación:

$$V_r = \left(1 + \frac{R_1}{R_2}\right)V_{ref} \quad (44)$$

donde  $R_1$  es una resistencia que realimenta negativamente al *OPAM* y variable para controlar el nivel de salida de  $V_r$ . Esto es importante ya que al trabajar luego con un *DSP* la señal debe tener un rango de nivel de continua entre 0 y 1 V. Las cuentas de los valores de los componentes se realizaron para que se cumpla esa restricción. Por lo tanto uno de los pots de la caja como indica la figura 4 es para controlar este valor. El amplificador inversor

toma la señal proveniente del resto del circuito, es decir la señal salida del filtro la invierte y amplifica, con la relación:

$$V_p = -\frac{R_3}{R_4}V_{in} \quad (45)$$

donde  $V_p$  es la salida y  $R_3$  la resistencia que realimenta el negativo que se es variable de forma tal de controlar la amplificación. Por lo tanto, el otro pote de la caja que muestra el esquema mostrado en la figura 4 la controla. Finalmente la última parte de esta etapa es un sumador que toma ambas señales ( $V_p$  y  $V_r$ ) y las suma así a la señal amplificada de la medición de presión se le agrega un valor de continua de forma independiente. Se pesan ambas señales, el nivel del continua y la señal amplificada, con resistencias previas al *OPAM* de forma tal de lograr a la salida del circuito una amplificación que satura en  $+12$  y  $-12$  V y un nivel de continua variable entre 0 y 1.

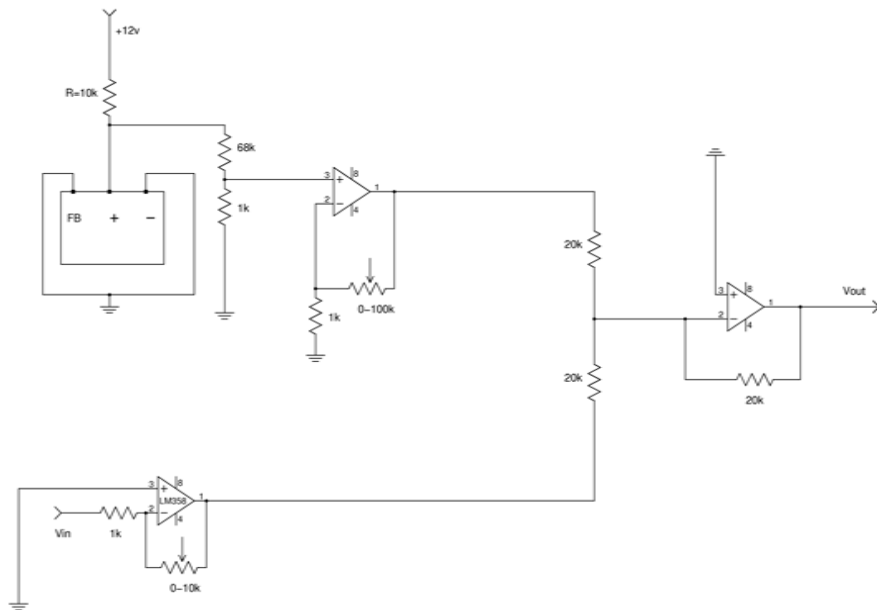
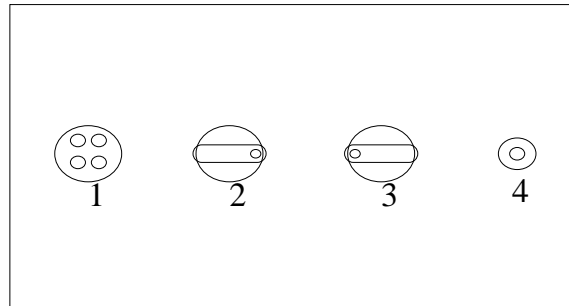


FIGURA 3. Esquema de la ultima etapa. En la parte de arriba a la izquierda se ve el circuito con el *LM385* y el amplificador seguidor. Abajo a la izquierda se muestra el esquema del amplificador inversor. A la derecha de la figura se muestra el sumador y la salida final del amplificador

---

### Vista Anterior



### Vista Posterior

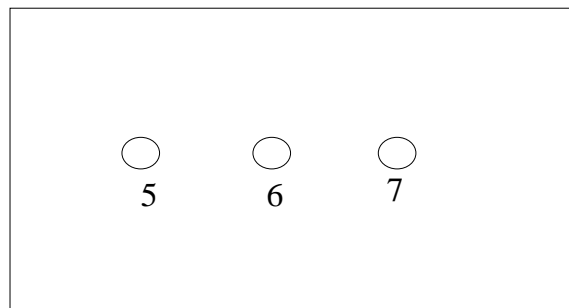


FIGURA 4. Esquema de la vista de la caja vista desde afuera. 1: entrada del conector 4 pines del *Fujikura*; 2: Potenciómetro que controla el nivel de continua; 3: Potenciómetro que controla la amplificación; 4: Salida de todas las etapas tipo BNC; en la parte posterior, 5:  $-12 V$ ; 6:  $0 V$ ; 7:  $+12 V$

## Encerar, pulir

Esta tesis pudo ser en un lugar en el que, antes que nada, se aprende a hacer ciencia. Se aprende que la ciencia no es física, ni biología, ni matemática, ni ingeniería: son preguntas y la mirada y la herramienta es la mezcla que haga falta para, si no encontrar las respuestas, al menos intentar comprender las preguntas.

Ese lugar funciona bien cuando es un equipo, que a mi me tocó vivir con Jorge y Jaco y Ana como “mayores”, que me mostraron dónde estaba todo y cómo había que hacer para que todo siga estando ahí y funcionando, y más que eso también: uno siempre aprende algún truco de los mayores, algo que guarda y se acuerda, de un guiño cuando estaba perdido, o asustado, o ansioso. También con los de más o menos mi generación, Lean y Mati y el espejo-desconcierto de encontrarnos ahí en horarios tan raros, como martes a las 11 de la noche o domingo a las 9 de la mañana.

Donde se armó ese grupo tan lindo con Nico, Gogui, Yoni, Flor, Marcos y algunos del otro lado del pasillo (los extranjerizados Martín y Belu, por ejemplo). Lo que sobra ahí es aguante, para trabajar y para pasar los momentos, los malos y los buenos.

Me resulta difícil imaginarme este lugar sin un Gabo que parece que todo el tiempo sabe por dónde hay que ir, y a la vez tiene la paciencia para que uno se dé cuenta a su tiempo, aunque tenga que equivocarse mil veces y alguna más hasta que salga bien. Si pudiera elegir una sola lección de las muchas que tuve la suerte de recibir en mi rato como alumno de Gabo, me quedaría con la que está en ese ejemplo, el de la vocación de mantener un espacio como el LSD donde se trabaja con comodidad e intensidad, de transmitir esa cultura de trabajo. Ojalá en mis ratos de doctorando “mayor” haya tomado bien la tarea de apuntalar ese proyecto desde donde me tocaba; ojalá haya en el futuro otras formas de aportar.

Agradezco la suerte de haber pasado por esta etapa en este laboratorio, con este equipo, bajo la dirección (y tantas veces el cuidado) de Gabo.





## Bibliografía

- [AAAM08] Ezequiel M. Arneodo, Leandro M. Alonso, Jorge A. Allende, and Gabriel B Mindlin. The dynamical origin of physiological instructions used in birdsong production. *Pramana journal of physics*, 70(6):1077, 2008.
- [AAGM09] Leandro M Alonso, Jorge A Allende, F Goller, and Gabriel B Mindlin. Low-dimensional dynamical model for the diversity of pressure patterns used in canary song. *Phys Rev E Stat Nonlin Soft Matter Phys*, 79(4 Pt 1):41929, April 2009.
- [AAM10] L M Alonso, J A Allende, and G B Mindlin. Dynamical origin of complex motor patterns. *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics*, pages 1–7, 2010.
- [AM08] Ana Amador and Gabriel B Mindlin. Beyond harmonic sounds in a simple model for birdsong production. *Chaos*, 18(May 2009):43123, 2008.
- [AM09] Ezequiel M Arneodo and Gabriel B Mindlin. Source-tract coupling in birdsong production. *Phys Rev E Stat Nonlin Soft Matter Phys*, 79(6):61921, June 2009.
- [AP90] D K Arrowsmith and C M Place. *An introduction to dynamical systems*. Cambridge University press, 1990.
- [APGM12] Ezequiel M Arneodo, Yonatan Sanz Perl, Franz Goller, and Gabriel B Mindlin. Model-based prosthetic vocal organ controlled by a freely behaving bird. *Plos comp. biology*, 2012. Manuscrito en preparación.
- [APM11] E.M. Ezequiel M Arneodo, Y.S. Yonatan Sanz Perl, and Gabriel B Mindlin. Acoustic signatures of sound source-tract coupling. *Physical Review E*, 83(4):041920, April 2011.
- [APMM] Ana Amador, Yonatan Sanz Perl, Daniel Margoliash, and Gabriel B Mindlin. Manuscrito en preparación.
- [aSGP99] Roderick a Suthers, Franz Goller, and Carolyn Pytte. The neuromuscular control of birdsong. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 354(1385):927–939, May 1999.
- [AT10] M F Assaneo and M A Trevisan. Computational model for vocal tract dynamics in a suboscine bird. *Phys. Rev. E*, 82(3):31906, September 2010.
- [BBM<sup>+</sup>03] C Beta, M Bertram, A S Mikhailov, H H Rotermund, and G Ertl. Controlling turbulence in a surface chemical reaction by time-delay autosynchronization. *Phys. Rev. E*, 67(4):46224, April 2003.

- 
- [BD00] Michael S Brainard and Allison J Doupe. Auditory feedback in learning and maintenance of vocal behaviour. *Nature Reviews Neuroscience*, 1:31–40, 2000.
- [BMA84] S W Bottjer, E A Miesner, and A P Arnold. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science*, 224(4651):901–903, 1984.
- [BNN96] M E Basham, E J Nordeen, and K W Nordeen. Blockade of NMDA receptors in the anterior forebrain impairs sensory acquisition in the zebra finch (*Poephila guttata*). *Neurobiology of learning and memory*, 66(3):295–304, 1996.
- [BPDLH97] E F Brittan-Powell, R J Dooling, O N Larsen, and J T Heaton. Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*). *Journal of the Acoustical Society of America*, 101(1):578–589, January 1997.
- [CB97] H J Chiel and R D Beer. The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20(12):553–557, 1997.
- [CFKS82] A B Cripps, A R Frey, L E Kinsler, and J V Sanders. Fundamentals of Acoustics, 1982.
- [CG04] Brenton G Cooper and Franz Goller. Partial muting leads to age-dependent modification of motor patterns underlying crystallized zebra finch song. *Journal of Neurobiology*, 61(3):317–332, 2004.
- [Cha02] Rulph Chassaing. *DSP Applications Using C and the TMS320C6x DSK*. John Wiley & Sons, Inc., New York, NY, USA, 2002.
- [CLC+03] Jose M Carmena, Mikhail A Lebedev, Roy E Crist, Joseph E O’Doherty, David M Santucci, Dragan F Dimitrov, Parag G Patil, Craig S Henriquez, and Miguel A L Nicolelis. Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS biology*, 1(2):E42, November 2003.
- [CMM+99] J K Chapin, K A Moxon, R S Markowitz, M A L Nicolelis, and Others. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature neuroscience*, 2:664–670, 1999.
- [CWR+08] A Cichocki, Y Washizawa, T Rutkowski, H Bakardjian, Anh-Huy Phan, Seungjin Choi, Hyekyoung Lee, Qibin Zhao, Liqing Zhang, and Yuanqing Li. Noninvasive BCIs: Multiway Signal-Processing Array Decompositions. *Computer*, 41(10):34–42, 2008.
- [CZPA09] A. Cichocki, R. Zdunek, A.H. Phan, and S. Amari. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. Wiley, West Sussex, UK, 2009.
- [DCM+06] S Darmanjian, G Cieslewski, S Morrison, B Dang, K Gugel, and J Principe. A Reconfigurable Neural Signal Processor (NSP) for

- 
- Brain Machine Interfaces. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 2502–2505. IEEE, 2006.
- [DK91] G. Dangelmayr and E. Knobloch. Hopf bifurcation with broken circular symmetry. *Nonlinearity*, 4:399, 1991.
- [EG08] Thomas Erneux and Johan Grasman. Limit-cycle oscillators subject to a delayed feedback. *Phys. Rev. E*, 78(2):26209, August 2008.
- [FKH04] Michale S Fee, Alexay A Kozhevnikov, and Richard H R Hahnloser. Neural Mechanisms of Vocal Sequence Generation in the Songbird. *New York Academy Sciences Annals*, 1016:153–170, June 2004.
- [FRS06] Neville H Fletcher, Tobias Riede, and Roderick A Suthers. Model for vocalization by a bird with distensible vocal cavity and open beak. *The Journal of the Acoustical Society of America*, 119(2):1005–1011, 2006.
- [FSPM98] Michale S Fee, Boris Shraiman, Bijan Pesaran, and Partha P Mitra. The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature*, 395(6697):67–71, 1998.
- [GC04] FRANZ Goller and BRENTON G Cooper. Peripheral Motor Dynamics of Song Production in the Zebra Finch. *Annals of the New York Academy of Sciences*, 1016(Behavioral Neurobiology of Birdsong):130–152, 2004.
- [GCM01] Tim Gardner, G Cecchi, and M Magnasco. Simple Motor Gestures for Birdsongs. *October*, (1):1–4, 2001.
- [GH83] John Guckenheimer and Philip Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag, New York Berlin Heidelberg Tokyo, 1983.
- [Gra87] Johan Grasman. *Asymptotic methods for relaxation oscillations and applications*. Springer-Verlag, New York, 1987.
- [GS95] Franz Goller and Roderick A Suthers. Implications for lateralization of bird song from unilateral gating of bilateral motor patterns. *Nature*, 373(6509):63–66, 1995.
- [GS96] F Goller and R A Suthers. Role of syringeal muscles in controlling the phonology of bird song. *J Neurophysiol*, 76(1):287–300, 1996.
- [HBTS95] Hanspeter Herzel, David Berry, Ingo Titze, and Ina Steinecke. Nonlinear dynamics of the voice: Signal analysis and biomechanical modeling. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):30–34, 1995.
- [HFH06] Haralambos Hatzikirou, W Tecumseh Fitch, and Hanspeter Herzel. Voice Instabilities due to Source-Tract Interactions. *Acta Acustica united with Acustica*, 92(3):468–475, 2006.
- [HFKF02] Richard H R Kozhevnikov Alexay A Hahnloser, Michale S Fee, Alexay A Kozhevnikov, and Michale S Fee. An ultra-sparse code underliethe generation of neural sequences in a songbird. *Nature*, 419(6902):65–70, March 2002.
-

- 
- [IF73] K Ishizaka and J Flanagan. *Synthesis of voiced sounds from a two-mass model of the vocal cords*. Dowden Hutchinson and Ross, 1973.
- [Keh05] Nasse Kehtarnavaz. *Real-Time Digital Signal Processing based on the TMS320C6000*. Elsevier, Oxford, UK, 2005.
- [KLL<sup>+</sup>09] Todd A. Kuiken, Guanglin Li, Blair A. Lock, Robert D. Lipschutz, Laura A. Miller, Kathy A. Stubblefield, and Kevin B. Englehart. Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms. *JAMA: The Journal of the American Medical Association*, 301(6):619–628, 2009.
- [LCiA04] Yuanqing Li, Andrzej Cichocki, and Shun ichi Amari. Analysis of sparse representation and blind source separation. *Neural Comput*, 16(6):1193–1234, June 2004.
- [LG99] O N Larsen and F Goller. Role of syringeal vibrations in bird vocalizations. *Proc Biol Sci.*, 266(1429):1609, 1999.
- [LGM02] Rodrigo Laje, Timothy J Gardner, and Gabriel B Mindlin. Neuromuscular control of vocalizations in birdsong: A model. *Phys. Rev. E*, 65(5):51921, May 2002.
- [LK99] Anthony Leonardo and Masakazu Konishi. Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature*, 399(6735):466–470, 1999.
- [LKCC07] H. Lee, Y.D. Kim, A. Cichocki, and S. Choi. Nonnegative tensor factorization for continuous eeg classification. *International journal of neural systems*, 17(4):305–318, 2007.
- [LM05] Rodrigo Laje and Gabriel B Mindlin. Modeling source-source and source-filter acoustic interaction in birdsong. *Phys. Rev. E*, 72(3):36218, September 2005.
- [LN06] M A Lebedev and M A L Nicolelis. Brain-machine interfaces: past, present and future. *TRENDS in Neurosciences*, 29(9):536–546, 2006.
- [LSZM08] R. Laje, D. Sciamarella, J. Zanella, and G.B. Mindlin. Bilateral source acoustic interaction in a syrinx model of an oscine bird. *Physical Review E*, 77(1):011912, 2008.
- [Mar86] D Margoliash. Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 6(6):1643–61, June 1986.
- [MB07] U. Meyer-Baese. *Digital signal processing with field programmable gate arrays*. Springer Verlag, 2007.
- [MGGS03] G. B Mindlin, T. J Gardner, F. Goller, and R. Suthers. Experimental support for a model of birdsong production. *Phys. Rev. E*, 68(4):41908, October 2003.
- [ML05] Gabriel Mindlin and Rodrigo Laje. *The Physics of Birdsong*. Springer, 2005.
- [Nic01] Miguel A L Nicolelis. Actions from thoughts. *Nature*, 409(6818):403–407, January 2001.
- [Now87] S Nowicki. Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature*,

- 
- 325(6099):53–55, January 1987.
- [OY97] Kazuo Okanoya and Ayako Yamaguchi. Adult bengalese finches require real-time auditory feedback to produce normal song syntax. *Journal of Neurobiology*, 33(4):343–356, 1997.
- [PAA<sup>+</sup>11] Yonatan Sanz Perl, Ezequiel M Arneodo, Ana Amador, Franz Goller, and Gabriel B Mindlin. Reconstruction of physiological instructions from Zebra finch song. *Phys. Rev. E*, 051909(5):1–8, November 2011.
- [PB81] J Perry and G A Bekey. EMG-force relationships in skeletal muscle. *Crit Rev Biomed Eng*, 7(1):1–22, 1981.
- [PHvH<sup>+</sup>94] X Pelorson, A Hirschberg, R R van Hassel, A P J Wijnands, and Y Auregan. Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model. *The Journal of the Acoustical Society of America*, 96(6):3416–3431, 1994.
- [PRK01] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization, a universal concept in nonlinear sciences*. Cambridge, University of Postdam, Germany, 2001.
- [PS00] L Parra and C Spence. Convolutional blind separation of non-stationary sources. *Speech and Audio Processing, IEEE Transactions on*, 8(3):320–327, May 2000.
- [RBF<sup>+</sup>97] Roger Randall, David Burggren, Warren French, Kathleen Fernald, and Russell Eckert. *Animal physiology : mechanisms and adaptation*, 1997.
- [RSFB06] Tobias Riede, Roderick A Suthers, Neville H Fletcher, and William E Blevins. Songbirds tune their vocal tract to the fundamental frequency of their song. *Proceedings of the National Academy of Sciences*, 103(14):5543–5548, April 2006.
- [SAGM08] JD D Sitt, A Amador, F Goller, and GB B Mindlin. Dynamical origin of spectrally rich vocalizations in birdsong. *Physical Review E*, 78(1):11905, July 2008.
- [SAGM10] JD Jacobo D Sitt, Ezequiel M EM Arneodo, Franz Goller, and Gabriel B Mindlin. Physiologically driven avian vocal synthesizer. *Phys. Rev. E*, 81(3):31927, March 2010.
- [SB06] J T Sakata and M S Brainard. Real-time contributions of auditory feedback to avian vocal motor control. *The Journal of neuroscience*, 26(38):9619–9628, 2006.
- [SCWM06] Andrew B Schwartz, X Tracy Cui, Douglas J Weber, and Daniel W Moran. Brain-controlled interfaces: movement restoration with neural prosthetics. *Neuron*, 52(1):205–20, October 2006.
- [SM02] Roderick A Suthers and Daniel Margoliash. Motor control of birdsong. *Current Opinion in Neurobiology*, 12(6):684–690, 2002.
- [Str00] Steven H Strogatz. *Nonlinear dynamics and chaos*. Perseus Publishing, Cambridge, Massachusetts, 2000.
- [Tex01a] Texas Instruments. *TLV320AIC23 Stereo Audio CODEC, 8 to 96 kHz, With Integrated Headphone Amplifier. Data Manual*,
-

- 
- 2001.
- [Tex01b] Texas Instruments. *TMS320C6000 Assembly Language Tools User's Guide*, 2001.
- [Tex04] Texas Instruments. *TMS320C6000 DSP Multichannel Buffered Serial Port (McBSP) Reference Guide*, 2004.
- [Tit88] Ingo R Titze. The physics of small-amplitude oscillation of the vocal folds. *The Journal of the Acoustical Society of America*, 83(4):1536–1552, 1988.
- [Tit00] I R Titze. *Principles of voice production*. National Center for Voice and Speech, 2000.
- [Tit08] Ingo R Titze. Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America*, 123(5):2733–2749, 2008.
- [TRP08] Ingo Titze, Tobias Riede, and Peter Popolo. Nonlinear source–filter coupling in phonation: Vocal exercises. *The Journal of the Acoustical Society of America*, 123(4):1902–1915, 2008.
- [WC72] Hugh R Wilson and Jack D Cowan. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophys. J.*, 12(1):1–24, 1972.
- [YM96] Albert C Yu and Daniel Margoliash. Temporal Hierarchical Control of Singing in Birds. *Science*, 273(5283):1871–1875, 1996.
- [ZM04] Philip Zeigler and Peter Marler. *Neuroscience of birdsong*. Cambridge University press., Cambridge, Massachusetts, 2004.
- [ZMP+05] D Zysman, J M Méndez, B Pando, J Aliaga, F Goller, and G B Mindlin. Synthesizing bird song. *Phys. Rev. E*, 72(5):51926, November 2005.
- [ZRS08] Sue Anne Zollinger, Tobias Riede, and Roderick A Suthers. Two-voice complexity from a single side of the syrinx in northern mockingbird *Mimus polyglottos* vocalizations. *The Journal of experimental biology*, 211(Pt 12):1978–91, June 2008.
- [ZS04] Sue Anne Zollinger and Roderick A Suthers. Motor mechanisms of a vocal mimic: implications for birdsong production. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1538):483–491, 2004.