

## Tesis Doctoral

# Biomimética vocal

Sitt, Jacobo Diego

2009

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:

Sitt, Jacobo Diego. (2009). Biomimética vocal. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Sitt, Jacobo Diego. "Biomimética vocal". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2009.

**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

## **Biomimética vocal**

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

**Médico Jacobo Diego Sitt**

**Directores de Tesis:** Dr. Gabriel B. Mindlin

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

Buenos Aires, 2009



## Biomimética vocal

### Resumen

El canto de las aves es un modelo para comportamiento vocal aprendido con gran similitud al proceso de aprendizaje y producción vocal en humanos. El canto es producido en el órgano vocal, la siringe, mediante la coordinada interacción de dos sistemas motores: los músculos que controlan la siringe y los músculos respiratorios. El canto está compuesto por elementos acústicos llamados sílabas, cada sílaba posee características propias. Mientras que algunas especies presentan únicamente sílabas tonales, otras como el Diamante Mandarín (*Taeniopygia guttata*) alternan sílabas tonales de alta frecuencia con sílabas de baja frecuencia que poseen gran contenido espectral.

En esta tesis estudiamos los mecanismos mediante los cuales la siringe puede contribuir a esa riqueza espectral. Se muestra que en el Diamante Mandarín, el gran contenido espectral presente en las sílabas de baja frecuencia, es una indicación de la dinámica presentada por su aparato vocal. Mediante el modelado teórico, esta dinámica es capturada en un modelo matemático de baja dimensión. La variación de contenido espectral no es el resultado de un control activo central de esta característica acústica, sino que emerge de la dinámica intrínseca de la siringe.

A partir de los modelos teóricos desarrollados de la biomecánica de la siringe, se diseñó un dispositivo biomimético, capaz de reemplazar el órgano vocal del ave. Se desarrolló una siringe electrónica capaz de generar canto mediante la transducción de instrucciones fisiológicas. La salida del dispositivo respeta las características acústicas del canto natural del ave. Este dispositivo electrónico integra digitalmente las ecuaciones de la dinámica de la siringe, constituyendo una solución robusta para generar vocalizaciones con control fisiológico en tiempo real.

Este trabajo ilustra la necesidad de explorar cuantitativamente la interacción entre el sistema nervioso y el sistema periférico que controla, de modo de entender el comportamiento emergente. Por otro lado, con el fin de generar dispositivos biomiméticos, se remarca la importancia de incorporar la física del problema en su desarrollo, logrando así soluciones más eficientes.

**Palabras clave:** Canto de aves. Control Motor. Dinámica no-lineal. Diamante Mandarín. Dispositivo biomimético

## Vocal biomimetic

### Abstract

Birdsong is a model system for learned vocal behaviour with remarkable parallels to human vocal development and sound production mechanisms. Song is generated in the birds vocal organ by the coordinated interaction of two motor systems the syrinx muscles and the respiratory muscles. Birdsong is composed by acoustic elements called syllable, each syllable has different acoustic features. While some species exhibit only tonal sounds, other species, like the Zebra Finch, alternate high frequency tonal notes and low frequency syllables with high spectral content.

In this thesis, we study the mechanisms of how the avian sound source might contribute to spectral richness. Here we show in the zebra finch (*Taeniopygia guttata*), that the broad range of upper harmonic content in different low frequency song elements is the fingerprint of the dynamics displayed by its vocal apparatus. Using theoretical modelling we captured this dynamic with a low dimensional dynamical model. The varying harmonic content of birdsong is not the result of the active neural control of the acoustic characteristics but emerges of the intrinsic dynamics of the sound source.

Using the developed theoretical models of biomechanics of the syrinx we also designed a biomimetical device capable of replacing the organ. We developed an electronic syrinx that generates song by transducing physiological instructions into acoustic output. The synthetic output is highly similar in acoustic features to the natural sounds. This electronic circuit digitally integrate the equations for the dynamics of the syrinx, constituting a robust solution suitable for generating synthetic song in real time.

This work illustrates the need to fully explore the deep interaction between a nervous system and the peripheral system it controls in order to understand the emerging behavior. Also, in order to develop biomimetic devices, we indicate the importance of incorporating the physics of the problem to design in order to obtain efficient solutions.

**Keywords:** Birdsong. Motor control. Nonlinear dynamics. Zebra Finch. Biomimetic device

Para Gise y Uri



---

## Agradecimientos

Es sorprendente cómo uno empieza a escribir esta página mucho antes de empezar el trabajo. Cinco años pueden ser un período importante en la vida de una persona, como también lo es lo que traemos en la mochila cuando arrancamos. En esa dirección quiero comenzar agradeciendo a mis padres, todo lo que pude lograr es gracias a ellos. Hoy me toca estar del otro lado del mostrador y valoro aún más todo lo que me dieron.

A Violeta y Rosa por los mimos desde chiquito. A Jack, se te extraña. A Pepe, más que un abuelo, un mejor amigo. Gracias por estar siempre a mi lado. A mis hermanas Denise, Jordana y Nicole, amigas, compañeras, compinches. Ayer, hoy, siempre, cerca o lejos nuestra relación será cada vez mas fuerte. No olviden que estaré siempre para ustedes.

A Silvia, Damian y Michelle gracias por hacerme uno mas de la familia. A Jorge, sentimos tu apoyo desde donde estés. A los amigos, los que llegan a nuestra vida para compartir. Los que estan cerca, los que se tuvieron que ir lejos. Gracias por acompañarme.

A la Universidad de Buenos Aires, la Facultad de Ciencias Exactas y Naturales, por darme una formación libre y de excelencia. A Jorge Aliaga por abrirme las puertas para iniciar mi doctorado, por con confiar en mi. A Mariano Sigman, por renovarme las ganas, por abrirme una puerta al futuro. Te doy mi palabra que la voy a dejar chiquita.

A todo el Laboratorio de Sistemas Dinámicos. Un gran grupo de gente. Jorge y Ana, gracias por compartir todo el camino. Marcos, tu aporte me llena de dicha. Zeque, un socio perfecto. Lean, gracias por la frescura. A los juniors, recuerden que el LSD es como de la familia y como tal hay que cuidarlo. Para todos ustedes, esta frase de Serrano: *“Y en el futuro espero, compañero, hermanos, ser un buen tipo, no traicionaros. Que el vértigo pase y que en vuestras ventanas luzca el sol cada mañana.”*

A mi director Gabriel, poca gente merece ser reconocida como un maestro que te da la vida. Gabriel, vos lo mereces. Más allá de la ciencia y de la admiración de estos años, me llevo haber aprendido a ser una mejor persona. Gracias por la paciencia.

Por último, y más importante, a quienes hacen mi regreso a casa todos los días. Quienes me han apoyado en todo momento, los días buenos y los no tanto. Por los momentos felices y por estar también muy juntos cuando no sale el sol. Gracias por formar parte de mi vida, por acompañarme en todas las aventuras. Los amo con todo el corazón.





# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Cuantificación, modelado matemático y biología . . . . .	1
1.2. Biomimética . . . . .	2
1.2.1. Biorobótica . . . . .	2
1.2.2. Neuroprostética . . . . .	3
1.3. Canto de Aves . . . . .	4
1.4. Esta tesis . . . . .	7
<b>2. Canto de aves</b>	<b>9</b>
2.1. Fisiología del canto de la aves . . . . .	9
2.1.1. Instrucciones centrales . . . . .	9
2.1.2. El aparato periférico . . . . .	10
2.2. Modelos físicos para el canto de aves . . . . .	16
2.2.1. Filtrado en el tracto vocal . . . . .	18
2.2.2. Validación experimental de los modelos . . . . .	21
<b>3. Diamante Mandarín</b>	<b>23</b>
3.1. Estudio experimental . . . . .	24
3.2. Un nuevo mecanismo dinámico . . . . .	29
3.3. Análisis dinámico . . . . .	35
3.4. Variables de control de las vocalizaciones . . . . .	37
<b>4. Biomimética vocal</b>	<b>41</b>
4.1. Modelo simplificado . . . . .	41
4.1.1. Análisis dinámico . . . . .	42
4.1.2. Escala temporal . . . . .	44
4.2. Siringe analógica . . . . .	49
4.3. Siringe digital . . . . .	54
4.3.1. ¿Qué es un DSP? . . . . .	54
4.4. Desarrollo de canto sintético . . . . .	56
4.5. Instrucciones fisiológicas conduciendo el DSP . . . . .	62
<b>5. Conclusiones</b>	<b>65</b>
<b>A. El Sonograma</b>	<b>71</b>
<b>Bibliografía</b>	<b>73</b>



# Introducción

---

## Índice

---

<b>1.1. Cuantificación, modelado matemático y biología . . . . .</b>	<b>1</b>
<b>1.2. Biomimética . . . . .</b>	<b>2</b>
1.2.1. Biorobótica . . . . .	2
1.2.2. Neuroprostética . . . . .	3
<b>1.3. Canto de Aves . . . . .</b>	<b>4</b>
<b>1.4. Esta tesis . . . . .</b>	<b>7</b>

---

## 1.1. Cuantificación, modelado matemático y biología

La aplicación de análisis cuantitativos en problemas biológicos tiene una larga historia, particularmente en la fisiología. Más recientemente, el vertiginoso aumento en la cantidad de datos experimentales biológicos ha hecho imperativo el trabajo interdisciplinario. El análisis cuantitativo y el modelado han ejercido una importante influencia en las siguientes áreas de la biología: organización celular (Keener and Sneyd, 2008), ecología y ecosistemas (May and Oster, 1976), evolución y diversidad (Hofbauer and Sigmund, 1988), organización y expresión del genoma (Smolen et al., 2000), crecimiento y desarrollo (Tomlin and Axelrod, 2007), sistema inmune y patógenos (Perelson, 2002), estructura y función molecular (Schlick, 2002), neurobiología y comportamiento (Izhikevich, 2006). Dentro de esta última área resulta de interés el uso del modelado matemático para explicar la emergencia de comportamiento complejo a partir de instrucciones neuronales de comando sencillas (Golubitsky et al., 1999).

Naturalmente, emergen barreras naturales de comunicación entre las disciplinas. Durante mucho tiempo la biología ha buscado desarrollar modelos verbales que describan los fenómenos observados, siendo el objetivo en la mayoría de los casos generar modelos cada vez más completos, con descripciones exhaustivas, enfatizando la diversidad de elementos que intervienen en los procesos asociados a la vida. Por otro lado los modelos matemáticos, son reduccionistas y tienden a menospreciar la diversidad reduciendo la descripción a su mínima expresión. El balance se logra cuando el modelo captura la esencia de las interacciones presentes, identifica los parámetros relevantes e interpreta sus resultados en el marco de las observaciones

experimentales.

Al analizar el aporte de los modelos matemáticos en la biología (y en cualquier otra disciplina) resulta importante responder a la siguiente pregunta: ¿Cuándo resulta útil el modelado matemático? Para responder a esta pregunta me atrevo a recurrir a la frase acuñada por Korzybski: “*El mapa no es el territorio*”. Es importante recordar que el modelo es una representación del fenómeno. La virtud del modelo será que su capacidad de síntesis elegante nos ayude a identificar y entender los mecanismos fundamentales que conforman el proceso estudiado.

En la práctica solo podremos medir la utilidad de un modelo matemático sabiendo cuánto podemos aprender del sistema biológico a partir del mismo. El objetivo ideal es poder hacer predicciones utilizando el modelo. Dichas predicciones deben ser lo suficientemente novedosas y deben invitarnos a volver al modelo experimental para testearlas.

## 1.2. Biomimética

La biomimética emerge de la intersección de múltiples disciplinas: física, biología, ingeniería y medicina. Su principal objetivo es el desarrollo de tecnología para reemplazar parte o la totalidad de un organismo biológico. Podemos separar esta empresa en corrientes principales, la biorobótica y la neuroprostética.

### 1.2.1. Biorobótica

La biorobótica tiene como objetivo el desarrollo de robots utilizando como inspiración a los organismos biológicos. Considerando el fin buscado en estos desarrollos, debemos dividir el área en dos. Llamaremos a la primera, robótica biológicamente inspirada. El foco principal de esta disciplina es la aplicación de ideas biológicas para solucionar problemas tecnológicos. La segunda, es el modelado biorrobótico. En este caso el objetivo es el uso de robots como una herramienta de modelado para responder preguntas biológicas.

En el primer caso el objetivo es plenamente tecnológico, la intención es obtener mejores robots. Al considerar lo eficaces y flexibles que son las soluciones encontradas por la naturaleza luego de millones de años de evolución, resulta prometedor incorporar en los robots aspectos de estas soluciones de modo de mejorar su rendimiento. Como ejemplos de este desarrollo se pueden presentar robots con locomoción inspirada en insectos y algoritmos de búsqueda basados en el comportamiento de fonotaxis en los grillos ([Webb and Consi, 2001](#)).

En el caso del modelado biorrobótico el objetivo es más académico. Del mismo modo que en cualquier otro modelo biológico, la intención es poder entender mejor las observaciones comportamentales y las mediciones neurobiológicas. Es creciente en la comunidad científica la idea de que el comportamiento animal debe ser comprendido no solo desde el sistema nervioso, sino desde la dinámica de su interacción con un cuerpo efector periférico y un medio ambiente ([Chiel and Beer, 1997](#)). La

ventaja del modelado biorobótico frente a los modelos computacionales, es que la interacción del modelo con un medio ambiente esta inherentemente incluida.

### 1.2.2. Neuroprostética

La neuroprostética es una disciplina en la intersección de la neurociencia y la ingeniería biomédica. Su principal interés es el desarrollo de prótesis comandadas por instrucciones neurales. Las prótesis neurales son una serie de dispositivos que pueden sustituir funciones motoras, sensoriales o cognitivas que pueden haber sido dañadas, ya sea por un traumatismo o por una enfermedad. El ejemplo fundacional de estos dispositivos es el de los implantes cocleares (Clark, 2003). Dicho dispositivo integra las acciones originalmente realizadas por una serie de órganos: filtrado por el conducto auditivo, acondicionamiento de la señal por el oído medio y análisis espectral por la cóclea. Resumiendo su funcionamiento, el dispositivo colecta los sonidos mediante un micrófono externo, procesa la señal usando un modelo de los órganos involucrados y estimula el nervio auditivo mediante un arreglo de microelectrodos implantados en el mismo. Otro ejemplo más reciente, pero no aún tan exitoso, es el de las prótesis visuales que permiten una recuperación parcial de la visión en no videntes (Zrenner, 2002).

Un caso particular de neuroprostética es el de las prótesis motoras para el control consciente del movimiento, llamado *brain machine interfaces* (Lebedev and Nicolelis, 2006). El principal objetivo es la restauración del control motor periférico, es decir, el control de actuadores periféricos (p.e., un brazo mecánico o una silla de ruedas). Para ello usan como entrada del dispositivo registros de la actividad nerviosa central o periférica. La actividad nerviosa central es registrada de forma invasiva o no invasiva. En el método invasivo el registro se realiza mediante microelectrodos implantados en cerebro del sujeto. De ese modo se busca medir la actividad en poblaciones neuronales específicas (Wessberg et al., 2000). En la alternativa no invasiva el registro de la actividad eléctrica cerebral se realiza externamente. Para ello se utilizan, por ejemplo, técnicas electroencefalográficas (EEG) (Millán et al., 2004). En el caso de actividad nerviosa periférica, el objetivo es registrar la actividad de nervios periféricos y de los músculos que ellos inervan. La técnica utilizada es llamada electromiograma (Basmajian, 1962; Junge) donde la actividad eléctrica es registrada ya sea mediante electrodos implantados en el músculo, o mediante electrodos de superficie.

### 1.3. Canto de Aves

El canto de las aves es uno de los sonidos más complejos y hermosos encontrados en la naturaleza. A lo largo de la historia ha inspirado a poetas y compositores. Es difícil encontrar ejemplos con mayor variedad y virtuosismo que el encontrado en su canto (Catchpole et al., 2003). Más allá de los motivos estéticos, esa inspiración es compartida por muchos hombres de ciencia. El principal motivo es que, en el 40 % de las 10000 especies de aves conocidas, dicho comportamiento complejo es aprendido.

El hecho de que el canto de las aves sea un comportamiento aprendido resulta interesante en gran parte porque es uno de los únicos modelos en el reino animal para el aprendizaje vocal en los humanos (Doupe and Kuhl, 1999). Dentro de los mamíferos el aprendizaje vocal está muy poco extendido. Solo los humanos, los cetáceos (ballenas y delfines), pocas especies de murciélagos y los elefantes (Poole et al., 2005) son capaces de aprender sus vocalizaciones. Por su parte, casi la mitad de las especies de aves conocidas presenta este comportamiento. Se trata de miles de especies que conforman el suborden Oscinos (también llamadas “aves canoras”), además de cientos de especies de loros (orden Psittaciformes) y colibríes (familia Trochillidae).

Es posible trazar un paralelo entre el proceso de aprendizaje del canto en las aves y el de adquisición del habla en humanos. En ambos casos, el aprendizaje puede dividirse en dos etapas. Una primera etapa “sensorial” y una segunda “sensorimotora”. En la primera etapa tanto aves como humanos escuchan las vocalizaciones de sus tutores para formar un templado interno. Peter Marler demostró que si se toman pichones en esta etapa y se los entrena a cantar bajo la tutela de adultos conspecíficos (de su misma especie), los mismos logran reproducir el canto de sus tutores. Por otro lado, si se los cria bajo aislamiento acústico o bajo la tutela de adultos de otra especie desarrollan canciones muy distintas a las típicas de su especie (Marler, 1970).

En la etapa siguiente los individuos generan subvocalizaciones (balbuceo en humanos y “sub-canción” en aves) que serán comparadas con el templado generado en la etapa anterior. Dicha etapa culmina con la obtención de vocalizaciones adultas. Konishi demostró que si se ensordecen pichones antes de que los mismos comiencen a realizar vocalizaciones estos desarrollaran un canto anormal (Konishi, 1965).

Una vez finalizada la segunda etapa, los humanos pueden seguir modificando sus vocalizaciones. Sin embargo, en las aves esta capacidad es variada. Existen especies, como el Diamante Mandarín (*Taeniopygia guttata*), que “cristalizan” su canto al finalizar la segunda etapa y no vuelven a incorporar nuevas sílabas. En el Diamante Mandarín esto sucede a los 90 días de vida. Por otro lado, otras especies, como el canario (*Serinus canaria*), pueden incorporar nuevas sílabas durante un período plástico que se repite cada temporada. Estos períodos están fuertemente controlados por la estacionalidad de los niveles hormonales.

Existe otro aspecto importante de la similitud entre el proceso vocal de aves y humanos. En ambos casos, la generación de sonido involucra mecanismos físicos

análogos. La génesis del sonido implica el paso de un flujo de aire a través de membranas: en humanos, las cuerdas vocales, ubicadas en la laringe; en aves, los labios siringeos, ubicados en la siringe. El flujo de aire transfiere energía a las membranas induciéndolas a vibrar. Las vibraciones modulan la presión en la base del tracto vocal generando así el sonido. El tracto vocal esta compuesto por estructuras análogas en aves y humanos. En ambos casos el tracto filtra, de manera controlada, el sonido antes de que el mismo se emita.

Las aves canoras constituyen probablemente el mejor modelo animal para estudiar los mecanismos implicados en un comportamiento complejo aprendido. Podemos enumerar las razones principales:

- el canto de las aves es un comportamiento repetitivo y en la mayoría de los casos muy sencillo de registrar. Puede ser grabado en forma automática utilizando simplemente un micrófono, una placa de adquisición y una PC. Desde la introducción por Thorpe de la técnica de sonogramas (Thorpe, 1958), el análisis espectrográfico de grabaciones de canto se ha simplificado notablemente (ver una explicación de la herramienta en el apéndice A);
- existen muchas especies de aves canoras que pueden estudiarse con facilidad en el laboratorio. Entre ellas, el Canario y el Diamante Mandarin se encuentran entre las más populares. De ese modo es posible introducir técnicas experimentales novedosas en un ambiente controlado;
- la descripción de los circuitos neuronales que controlan la producción y el aprendizaje del canto permite intentar identificar su función específica y la interacción entre ellos. Los circuitos están compuestos por un pequeño número de núcleos interconectados (Margoliash, 1997; Spiro et al., 1999). Diversos laboratorios han desarrollado técnicas que permiten registrar la actividad en estos núcleos mientras el ave se encuentra cantando;
- en el marco del estudio de los núcleos neuronales que controlan el canto de las aves. Fernando Nottebohm hizo, a mi entender, el descubrimiento más importante de la neurociencia moderna, el descubrimiento de neurogénesis en el cerebro adulto de vertebrados (Goldman and Nottebohm, 1983). Este descubrimiento dio por tierra una de las teorías más arraigadas en neurociencia: que el sistema nervioso adulto tiene un número fijo de neuronas y que las mismas no pueden ser reemplazadas por neuronas nuevas. La neurogénesis abre la oportunidad de estudiar el rol que cumplen estas nuevas neuronas en procesos como el aprendizaje y la memoria (Kempermann et al., 2004);



- el canto de aves ofrece una excelente oportunidad de estudiar cómo el control neuromuscular coordinado de estructuras periféricas, contribuye a la emergencia de comportamiento. Considerando la complejidad del comportamiento observado, el estudio de las instrucciones de control nos permite intentar responder las siguientes preguntas: ¿cuánta complejidad está presente en las instrucciones? y ¿cuánta emerge de la interacción entre las instrucciones centrales, la fisiología de las estructuras periféricas y el medio ambiente? (Chiel and Beer, 1997);
- el proceso de aprendizaje implica una integración sensorimotora. Esto brinda la oportunidad de estudiar el feedback auditivo de sus propias vocalizaciones que el ave modifica ajustando las instrucciones de control motor;
- por último, el constante crecimiento de las técnicas moleculares nos permite preguntarnos por los genes que determinan la capacidad de una especie para poseer un comportamiento vocal aprendido. Muy recientemente se ha identificado un pequeño grupo de genes, cuyos circuitos de activación están relacionados con el aprendizaje vocal tanto en humanos como en aves (White, 2009).

## 1.4. Esta tesis

Esta tesis está dedicada a continuar el estudio de los mecanismos físicos en la fonación en las aves y analizar las aplicaciones del conocimiento aprendido al desarrollo de dispositivos biomiméticos. Para ello utilizamos modelos matemáticos para describir el comportamiento del órgano vocal. Usando herramientas de dinámica no lineal, analizamos estos modelos y extraemos características paradigmáticas. Utilizando estos modelos se implementa un dispositivo neuroprotético vocal. El equipamiento es comandado en tiempo real por instrucciones fisiológicas.

La tesis se encuentra organizada de la siguiente manera:

### **Capítulo 2: Revisión del estudio de canto de aves**

En un primer paso se analiza el aporte de la biología al estudio del canto de las aves. Luego se revisa el estado del arte en el modelado físico de la generación del canto en el órgano vocal.

### **Capítulo 3: Diamante Mandarín y contenido espectral**

El Diamante Mandarín es la especie más utilizada en el estudio del canto de las aves. Sin embargo, los mecanismos físicos que determinan el complejo contenido espectral de sus vocalizaciones no ha sido completamente estudiado. En este capítulo se propone un nuevo modelo físico para la fonación. Se identifica el régimen dinámico y los parámetros relevantes para la síntesis de sonidos en el modelo. Finalmente, se presenta la hipótesis de que gran parte de la complejidad espectral encontrada es una indicación de las propiedades no-lineales de la siringe.

### **Capítulo 4: Desarrollo de dispositivo biomimético vocal**

En este capítulo, se presenta la implementación de un equipamiento biomimético vocal para el Diamante Mandarín. Para ello se utiliza el modelo desarrollado en el capítulo anterior. Se identifican las variables fisiológicas necesarias para poder comandar el modelo. Las señales fisiológicas se adaptan adecuadamente para introducir las en el dispositivo. De ese modo se logra comandar el mismo en tiempo real mediante instrucciones fisiológicas registradas en el ave.

### **Conclusiones**

En este capítulo analizaremos brevemente los hallazgos y aportes realizados. También se discuten nuevas líneas de trabajo.

# Canto de aves

---

## Índice

---

<b>2.1. Fisiología del canto de la aves . . . . .</b>	<b>9</b>
2.1.1. Instrucciones centrales . . . . .	9
2.1.2. El aparato periférico . . . . .	10
<b>2.2. Modelos físicos para el canto de aves . . . . .</b>	<b>16</b>
2.2.1. Filtrado en el tracto vocal . . . . .	18
2.2.2. Validación experimental de los modelos . . . . .	21

---

En este capítulo revisaremos brevemente el estado del arte en el estudio del canto de las aves. Comenzaremos con el estudio de las instrucciones neuronales que lo controlan, los órganos periféricos que lo ejecutan y la interacción entre ambos. Luego estudiaremos el aporte desde la física para comprender los mecanismos presentes en la generación de canto en el órgano vocal.

## 2.1. Fisiología del canto de la aves

El estudio de la fisiología del canto de las aves puede dividirse en dos áreas: la de las instrucciones centrales que controlan la ejecución del proceso y la respuesta de los aparatos periféricos responsables de la mecánica del canto. El comportamiento emerge de la interacción entre estas dos áreas. Recientemente el aporte de modelos integrativos ha generado evidencia en favor que mucha de la complejidad del comportamiento emerge de la interacción y no de los procesos centrales y periféricos en forma individual (Chiel and Beer, 1997).

### 2.1.1. Instrucciones centrales

Los pasos fundacionales en el estudio moderno del canto de las aves comenzaron con el descubrimiento de núcleos neuronales específicos en el cerebro necesarios para la producción de canto aprendido, a partir de los déficits en el comportamiento vocal producidos al lesionar distintas áreas del cerebro (Nottebohm et al., 1976). El estudio del funcionamiento de dichos núcleos continuó con los primeros experimentos de electrofisiología realizados durante el canto (McCasland and Konishi, 1981). Más recientemente, el uso de técnicas moleculares y comportamentales han sido útiles para descubrir la existencia de los mismos grupos neuronales en colibríes (Jarvis et

al., 2000).

Mucho se ha aprendido respecto al rol que cumplen cada uno de estos núcleos en la síntesis del canto. Actualmente, la comunidad que estudia el canto de aves ha consensuado una clasificación de estos núcleos según su funcionalidad: un primer grupo está compuesto por los núcleos necesarios para vocalizar, formando la *vía motora o vía posterior* y otro grupo de núcleos involucrado en el aprendizaje y mantenimiento de las vocalizaciones. Este grupo es llamado *AFP* (por sus siglas en inglés *Anterior Forebrain Pathway*).

Básicamente, la vía motora está compuesta por dos núcleos: HVC (por *hypers-triatum ventrale pars caudale* aunque posteriormente rebautizado *High Vocal Center*) y el RA (por *robustus nucleus of the archistriatum*). El HVC envía instrucciones al RA, quien a su vez envía instrucciones a múltiples núcleos: el núcleo traqueosiríngeo (o nXIIIts), que inerva los músculos siringeos; y una serie de núcleos respiratorios (PAm y RA<sub>m</sub>) responsables del control de dicho proceso (ver un esquema de los núcleos en la Figura 2.1). HVC y RA están vinculados al programa del canto, organizados jerárquicamente: microestimulaciones en HVC llevan a una interrupción y reinicio general del canto mientras que el mismo procedimiento en RA lleva a alteraciones transitorias del canto (Yu and Margoliash, 1996). De este modo, HVC tendría una función de planificación general del canto y RA la ejecución de dicho plan.

La vía AFP conecta indirectamente HVC y RA como se muestra en la Figura 2.1. Lesiones en los núcleos que conforman dicha vía no alteran el canto en forma inmediata pero afectan la capacidad de aprendizaje de las aves (Bottjer et al., 1984). La vía esta formada por tres núcleos: el área X, DLM (*medial nucleus of the dorsala-teral thalamus*) e IMAN (*lateral magnocellular nucleus of the anterior neostatum*). Si bien existen observaciones parciales que reportan el funcionamiento de esta vía, aún no ha sido posible componer una teoría para su comportamiento global.

### 2.1.2. El aparato periférico

El estudio del aparato periférico tiene como fin entender el proceso físico de producción del canto. Conocer el proceso de funcionamiento, sus capacidades y limitaciones, es vital para entender el proceso vocal.

El órgano vocal de las aves es la siringe. Su ubicación y morfología varía entre especies (King, 1989). Pueden clasificarse en las que tienen una única fuente sonora y las que tienen dos fuentes sonoras. En las primeras la siringe está ubicada en la tráquea y está presente en loros y palomas. La segundas pueden dividirse en dos tipos: las bronquiales (como las de pingüinos y lechuzas) y las traqueo-bronquiales (presentes en todo el suborden oscinos) (Suthers, 2001). Esta última se encuentra ubicada en el lugar donde confluyen los dos bronquios primarios formando la traquea. Esta clase de siringe resulta de nuestro interés ya que es el caso de estudio en esta tesis. El Diamante Mandarín presenta una siringe de este tipo. Se presenta un esquema de la siringe en la Figura 2.2.

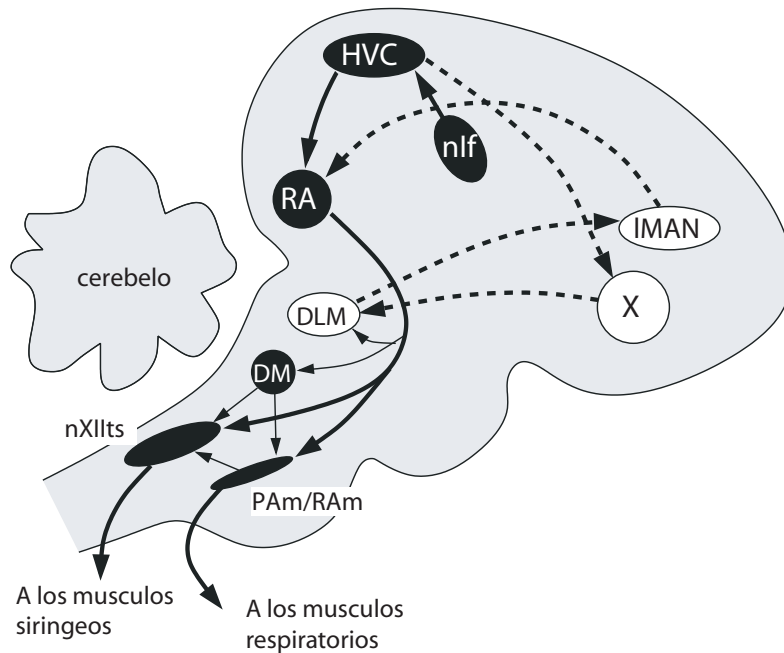


Figura 2.1: Esquema de los núcleos que conforman el sistema de control del canto. En la vía motora los núcleos se indican con óvalos sólidos y las conexiones con flechas continuas. Para la vía anterior (AFP) los núcleos se indican con óvalos huecos y las conexiones con flechar rayadas. Modificado de [Spiro et al. \(1999\)](#).

Actualmente existe acuerdo en la comunidad científica respecto al mecanismo de producción de vocalizaciones en la siringe. Estas se generan por oscilaciones de estructuras blandas en las paredes internas de la siringe, inducidas por transferencia de energía del flujo de aire impulsado por los sacos aéreos durante las vocalizaciones. Este proceso es análogo al que ocurre en las cuerdas vocales de los humanos ([Titze; Gardner et al., 2001](#)). En el caso de la siringe traqueo-bronquial observaciones endoscópicas indican que las estructuras que oscilan son los labios laterales y mediales ubicados en el extremo cefálico de los bronquios (LL y ML en la Figura 2.2) ([Larsen and Goller, 1999](#)).

Como se puede ver en la Figura 2.2 existe una gran cantidad de músculos que actúan sobre la siringe: los intrínsecos (con ambas inserciones dentro de la siringe) y los extrínsecos (con una inserción dentro y otra fuera de la siringe). La función de los músculos se estudia principalmente mediante el uso de la técnica de electromiografía. En dicha técnica se registra la actividad mecánica de los músculos en forma indirecta mediante la actividad eléctrica en los mismos ([Randall et al., 2001](#)). Si bien la función específica de cada músculo y la interacción entre ellos aún no está del todo develada, se puede relacionar la actividad de los músculos ventrales con la frecuencia de oscilación de los labios y la actividad de los músculos dorsales con la habilitación de las oscilaciones o gating.

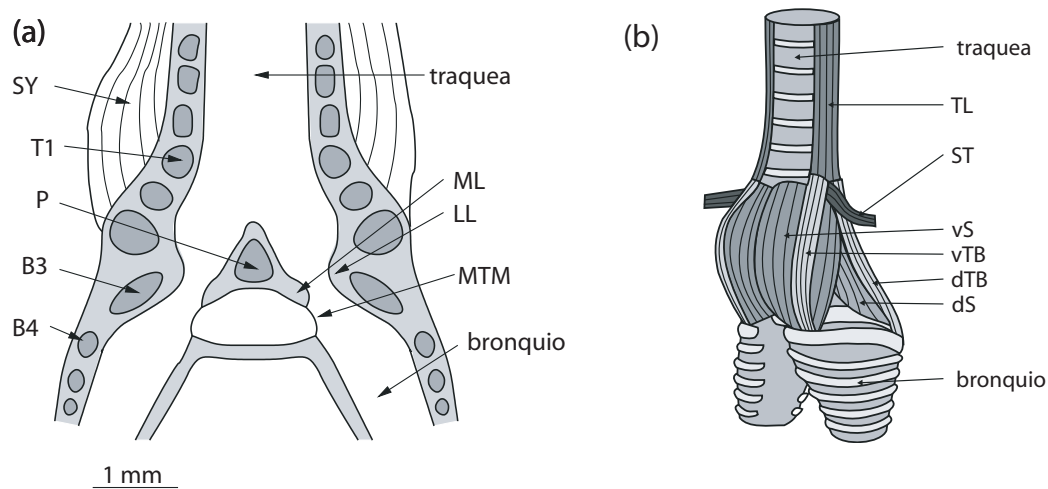


Figura 2.2: Esquema de la siringe traqueo-bronquial. (a) Corte coronal de la siringe, en el mismo se observan: LL y ML, labios laterales y mediales; MTM, membrana timpaniforme medial; T1, primer anillo traqueal; B3 y B4, tercer y cuarto anillo bronquial; P, pesulo; SY, músculo siringeo (b) Vista ventro-lateral de la siringe, se muestran los músculos involucrados en la producción del canto las abreviaturas son: vS, *synrigealis ventralis*; dS, *synrigealis dorsalis*; TL, *tracheolateralis*; dTB, *tracheobronchialis dorsalis*; ST, *sternotrachealis*; vTB, *tracheobronchialis ventralis*. Modificado de Mindlin and Laje (2005) con autorización del autor

La actividad de los músculos dorsales controla la ubicación relativa de los labios laterales y mediales. Cuando el ave respira sin emitir canto estos músculos se encuentran relajados; los labios se alejan entre sí impidiendo la oscilación. Cuando el ave se prepara para vocalizar, los músculos se contraen y los labios se acercan, generando una obstrucción parcial del lumen bronquial. Cuando la presión supera un umbral los labios comienzan a oscilar, generando así el sonido (Goller and Larsen, 1997; Larsen and Goller, 1999). Si la actividad de los músculos es aun mayor, los labios colisionan entre sí cerrando el lumen, bloqueando de ese modo las oscilaciones en esa fuente de la siringe (Larsen and Goller, 2002).

Por otro lado los músculos ventrales regularían la frecuencia de oscilación de los labios. Se ha encontrado en canarios y cardenales una correlación entre la actividad de los músculos ventrales y la frecuencia fundamental de oscilación (Suthers, 2001). El mecanismo de control resultaría de la modificación en la tensión de los labios.

Las propiedades mecánicas de los músculos de la siringe merecen una mención especial. Resultados recientes los ubican como uno de los ejemplos de respuesta muscular más rápidos del reino animal (Elemans et al., 2008). La velocidad de contracción de estos músculos “súper rápidos” se corresponde con la estructura temporal encontrada dentro del canto.

Para producir canto, el ave debe no solo controlar el sistema motor que acciona los músculos de la siringe sino que también en forma coordinada debe controlar un conjunto de músculos que regulan el sistema respiratorio (Suthers et al., 1999). El

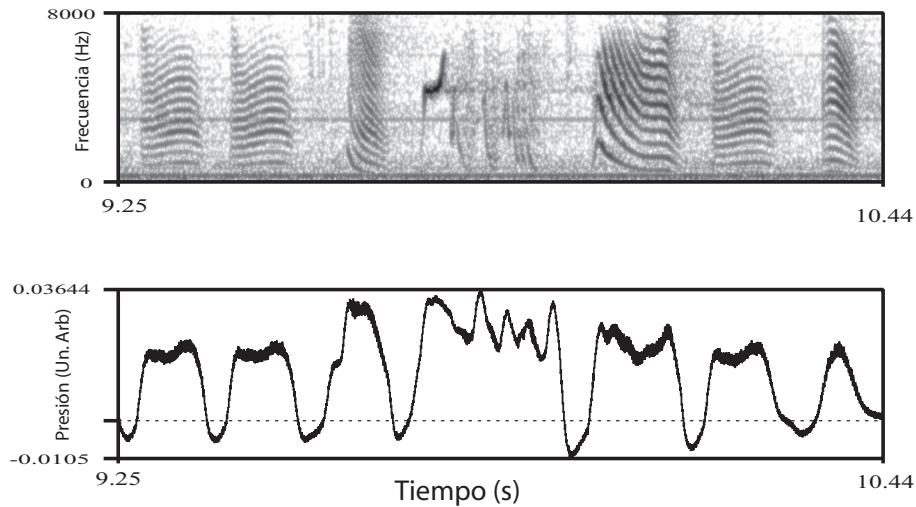


Figura 2.3: Registro típico de canto de Diamante Mandarin. En el panel superior se observa el sonograma de un canto típico y en el panel inferior la presión en los sacos aéreos en el momento de la vocalización.

sistema respiratorio aviar es más complejo que el de los humanos. Los pulmones son rígidos y el control de la ventilación se produce por unas estructuras flexibles llamadas sacos aéreos. El control de la presión en los sacos aéreos está dado por los músculos respiratorios. La gran mayoría de los sonidos generados en por las aves se produce durante la fase espiratoria del ciclo respiratorio. Durante la generación de canto la presión en los sacos aéreos aumenta alrededor de un orden de magnitud respecto de la respiración normal (Hartley and Suthers, 1989). Los periodos de silencio entre las sílabas se corresponden con inspiraciones cortas, necesarias para recargar el aire espirado. Los patrones respiratorios determinan en gran medida la estructura temporal del canto. En la Figura 2.3 se muestra simultáneamente el sonograma de un registro típico de vocalización de Diamante Mandarin y la presión en sus sacos aéreos.

El sonido que emite la siringe viaja a través de una serie de estructuras (traquea, glotis, cavidad orofaríngea y pico) que conforman lo que llamaremos el tracto vocal. El hecho de pasar por dicho tracto produce una amplificación de ciertas bandas de frecuencias de sonido emitido por la siringe y la atenuación de otras. El tracto vocal posee musculatura que potencialmente permitiría modificar su configuración. De ese modo el ave tiene mecanismos para lograr un control activo del filtrado (Fletcher and Tarnopolsky, 1999). Analizaremos con mayor detalle el proceso de filtrado en la sección 2.2.1.

### 2.1.2.1. Fisiología de la periferia en Diamante Mandarin

En su revisión de 2004 Goller y Cooper hicieron una síntesis de la dinámica motora periférica en Diamante Mandarin (Goller and Cooper, 2004). El canto de estas



aves consiste de un repertorio de 3 a 8 sílabas las cuales son repetidas un número variable de veces (Williams, 2004)(ver Figura 2.3). Las vocalizaciones pueden ser clasificadas en dos grupos: las de baja frecuencia fundamental (400-800  $Hz$ ) con gran riqueza espectral y las de alta frecuencia fundamental (3-7  $kHz$ ) con comportamiento espectral tonal. Analizaremos en mucho mayor detalle la relación entre frecuencia fundamental y contenido espectral en el próximo capítulo.

En las vocalizaciones de baja frecuencia, la dinámica de oscilaciones de los labios siringeos en Diamante Mandarin tendrían un patrón distinto al exhibido por otras aves oscinas. Recientemente se reportó que dichas vocalizaciones se producen mediante una dinámica de oscilaciones pulsátiles de los labios (Jensen et al., 2007). Dicha dinámica ha sido también reportada en humanos (Blomgren et al., 1998) siendo responsable del registro vocal más bajo, llamado *vocal fry*. Una de las características más importantes de este registro es la independencia de escala, es decir, la decorrelación entre las dimensiones físicas del órgano fonador con la frecuencia de oscilación. Esta independencia de escala explicaría un fenómeno que por mucho tiempo intrigó a los investigadores; cómo un ave tan pequeña puede emitir vocalizaciones de tan baja frecuencia.

El aporte relativo de cada fuente de la siringe en el canto fue cuantificado registrando el flujo aéreo en cada bronquio. Para ello se insertaron termistores configurados como sensores de flujo en el lumen de los bronquios (Goller and Daley, 2001). Estas mediciones reportaron una serie de resultados importantes, el primero es que los sonidos emitidos por la fuente derecha tienen una frecuencia media mayor a los emitidos por la fuente izquierda. También se reportó que las vocalizaciones de alta frecuencia son emitidas en forma exclusiva por la fuente derecha. La clasificación de las vocalizaciones de baja frecuencia es más compleja. En todos los casos de vocalizaciones de baja frecuencia la fuente derecha se encuentra activa. Simultáneamente, la fuente izquierda se encuentra siempre activa en algunas sílabas y en parte de la duración del resto de las sílabas.

Los patrones de activación de los músculos siringeos en Diamante Mandarin difieren de lo observado en otras especies. En su trabajo de principios de los '90 Vicario reportó la ausencia de actividad ventral en las vocalizaciones de baja frecuencia (Vicario, 1991). Este resultado es contradictorio con lo observado en otras aves oscinas (Goller and Suthers, 1996). Por ello se requiere estudiar con más detalle cómo en esta especie las instrucciones centrales controlan las características acústicas del canto. Aun más, en esta especie ha sido sugerido que las propiedades no lineales de la siringe contribuyen a la complejidad del canto sin requerir instrucciones neuronales complejas (Fee et al., 1998; Tchernichovski et al., 2001).

El control motor del tracto superior en Diamante Mandarin fue inicialmente estudiado por Daley y Goller. Los datos experimentales indicarían que los cambios en longitud de la tráquea son despreciables y no modificarían las características del filtrado (Daley and Goller, 2004). Por otro lado, existe evidencia de que la apertura del pico modificaría las propiedades del tracto generando un énfasis en los armónicos

superiores. Estas conclusiones son consistentes con una hipótesis de no interacción entre la fuente y el filtro (Goller et al., 2004).

Yendo más allá de esta hipótesis, ciertas características acústicas del canto pueden interpretarse como manifestaciones de fenómenos no lineales característicos de los sistemas no lineales realimentados (Arneodo and Mindlin, 2009). Resultados experimentales muy recientemente obtenidos en el Laboratorio de Sistemas Dinámicos sugieren que algunas sílabas la apertura del pico puede ser utilizada como un mecanismo de control de la fonación (Sanz et al., 2009). Es por ello que a futuro, será necesario re-evaluar la importancia del control motor del tracto superior en esta ave.

## 2.2. Modelos físicos para el canto de aves

La consideración de todas las fuentes potenciales para la emergencia de complejidad en el canto de las aves dificulta extraordinariamente el tratamiento teórico global del problema. Sin embargo, trabajo teórico reciente sugiere un modo de distinguir qué parte de la complejidad del canto emerge de propiedades no lineales intrínsecas de su órgano fonador, de la interacción con las instrucciones centrales y el medio ambiente. En esta sección haremos una revisión de los modelos físicos que explican el canto de las aves.

Recordemos que las estructuras responsables de la generación de sonidos son los labios laterales y mediales, ubicados en el espacio luminal de los bronquios. Para que dichas estructuras se pongan a oscilar es necesario que se dispongan en una posición relativa adecuada y que exista un flujo aéreo entre ellas.

En el marco del estudio de oscilaciones en las cuerdas vocales de humanos, Titze propuso uno de los modelos físicos más sencillos para explicar la aparición de oscilaciones auto-sostenidas (Titze, 1988). En dicho modelo se descompone el movimiento de los labios en dos movimientos simples. El primero corresponde a los labios acercándose y alejándose entre sí. El segundo movimiento corresponde a una onda de propagación ascendente. (Ver un esquema de dichos movimientos en los paneles (b1) y (b2) de la Figura 2.4). Veremos que la correcta combinación de dichos movimientos es necesaria para la aparición de oscilaciones.

Si consideramos que el ave se encuentra espirando en el momento de la vocalización, la presión por debajo de los labios es mayor que la atmosférica y la presión por encima de los labios es prácticamente igual a la atmosférica. La combinación que permite la aparición de oscilaciones es tal que los labios tengan un perfil divergente cuando se acercan y más plano cuando se alejan. De ese modo la presión ejercida sobre los labios será mayor en la apertura y menor en la clausura. Este efecto determina que el sistema gane energía en cada movimiento. En el panel (c) la Figura 2.4 se esquematiza la secuencia de dichos movimientos.

Aproximando la forma de los labios como plana (ver Figura 2.4), la dinámica de los labios puede modelarse como,

$$a_1 = a_{01} + x + \tau \frac{dx}{dt} \quad (2.1)$$

$$a_2 = a_{02} + x - \tau \frac{dx}{dt}, \quad (2.2)$$

donde  $a_1$  y  $a_2$  son las separaciones medias de los labios correspondientes a los extremos superior e inferior respectivamente (ver Figura 2.4). El parámetro  $\tau$ , corresponde al tiempo que le toma a la onda labial propagarse la mitad de la dimensión vertical del labio,  $a_{01}$  y  $a_{02}$  a las separaciones medias en el estado de reposo de los extremos de los labios (Gardner et al., 2001).

Siguiendo la metodología usada por Titze (Titze, 1988), se puede calcular la presión media en los labios como,

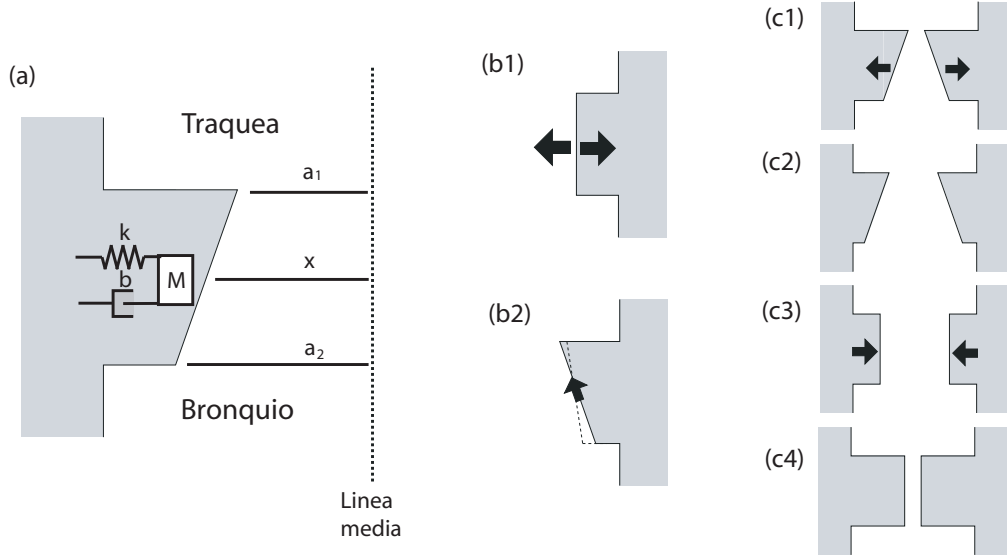


Figura 2.4: Esquema de parámetros del modelo y dinámica de movimiento de los labios siringeos. En (a) se indican parámetros del modelo. En el paneles (b1) se muestra el movimiento lateral y en (b2) la onda ascendente, para un solo lado de la siringe. En (c) se muestra la secuencia de combinación de movimientos (b1) y (b2) necesarios para generar oscilaciones auto-inducidas.

$$p_{media} = p_{sub} \left(1 - \frac{a_2}{a_1}\right), \quad (2.3)$$

donde  $p_{sub}$  corresponde a la presión subsiringea.

Para escribir la ecuación que rige la dinámica del movimiento para la posición media del labio ( $x$ ) consideraremos las fuerzas presentes. Utilizando la ley de Newton la misma queda:

$$\frac{dx}{dt} = y \quad (2.4)$$

$$m \frac{dy}{dt} = -kx - by - cx^2y - g_0 + a_{lab} p_{sub} \frac{\Delta a + 2\tau y}{a_{01} + x + \tau y}. \quad (2.5)$$

Donde  $m$  corresponde a la masa del labio. El primer término corresponde a la fuerza de restitución de los labios. En este término el parámetro  $k$  esta asociado la tensión. El segundo término representa la disipación del sistema modelada a través del parámetro de disipación  $b$ . El tercer término, también una disipación no lineal, se torna relevante cuando  $x$  toma valores grandes, correspondiendo a alejamientos grandes respecto la posición de equilibrio. De ese modo, esta fuerza de disipación no lineal sirve para modelar las colisiones entre los labios o con las paredes continentales. De ese modo se introduce en el modelo las restricciones naturales al movimiento. El termino  $g_0$  representa una fuerza independiente a la dinámica de los labios la cual modela el gating muscular activo (Laje et al., 2002).

Finalmente, el último término describe la fuerza actuando debido a la presión inter-labial. Dicho termino se calcula reemplazando las ecuaciones 2.1 y 2.2 en 2.3. El parámetro  $\Delta a = a_1 - a_2$  define la posición prefonatoria. La fuerza se calcula multiplicando la presión por el área labial representada por  $a_{lab}$ .

Este modelo tiene la característica que, aun siendo una representación de baja dimensión del problema, sus parámetros tienen representación biológica directa. El control del sonido generado por el modelo se logra mediante la modulación lenta de los parámetros  $k$ ,  $g_0$  y  $p_{sub}$ . El estudio del comportamiento del sistema para dos de estos parámetros, presión y tensión, permite una descripción matemática del mecanismo de oscilación. Por ejemplo, una de las predicciones es que las oscilaciones arrancan cuando la presión supera un valor umbral. El mecanismo dinámico de encendido de dichas oscilaciones es mediante una bifurcación Hopf. Las oscilaciones que nacen se caracterizan por ser tonales puras, tener amplitud cero y frecuencia bien definida. (Strogatz, 2001). Para valores mayores de presión, aumentará la amplitud de las oscilaciones y el contenido espectral de las mismas, pero la frecuencia de oscilación se mantendrá prácticamente constante. La frecuencia de oscilación aumentara con el valor del parámetro  $k$ .

La salida de este modelo, adecuadamente filtrada como veremos más adelante, ha sido usado con éxito para sintetizar canto de especies que caracterizan por vocalizaciones tonales puras (por ejemplo Canarios, Chingolos y Cardenales). Las sílabas son generadas recorriendo elipses en el espacio  $(p_{sub}, k)$

### 2.2.1. Filtrado en el tracto vocal

Las oscilaciones generadas en la siringe introducirán fluctuaciones de presión en la base de la tráquea. Dichas oscilaciones deberán recorrer una serie de estructuras semi-rígidas antes de llegar a la atmósfera. Estas estructuras conforman lo que llamamos el tracto vocal o tracto superior.

La función del tracto vocal varía entre especies. Las especies que tienen vocalizaciones tonales sintonizarían el tracto de modo de enfatizar la frecuencia fundamental (Riede et al., 2006; Nowicki, 1987). Para las especies con un canto espectralmente rico, como es el caso del Diamante Mandarín presentado previamente, la función del tracto vocal está menos estudiada.

Existen dos estrategias para acoplar la dinámica de la fuente con el proceso de filtrado producido en el tracto vocal. En la primera no se asumen interacciones entre la fuente y el filtro. Es decir, el tracto realiza un filtrado pasivo de la salida de la fuente sin afectar su dinámica. En este caso ambos sistemas se computan en forma desacoplada. Se llama a esta hipótesis “ hipótesis de separación fuente-filtro ” (ver Figura 2.5). La otra hipótesis considera la interacción de la fuente con el filtro. Para ello, se debe incorporar en el modelo de la fuente las fluctuaciones de presión debidas a reflexiones en el tubo. En este caso los sistemas deben computarse en forma simultánea. Se llama a esta hipótesis “ hipótesis de interacción fuente-filtro ”. En el marco de la misma se han reportado fenómenos complejos como duplicaciones de periodo o saltos de frecuencia (Laje and Mindlin, 2005; Arneodo and Mindlin,

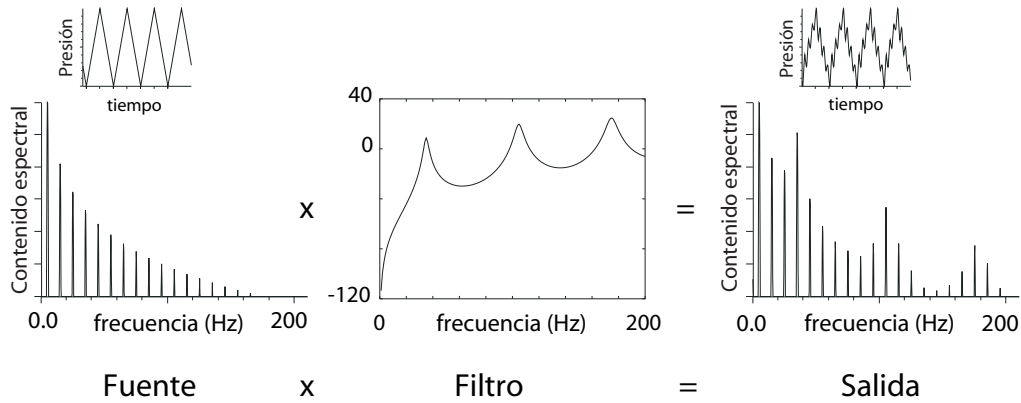


Figura 2.5: Filtrado pasivo de una señal. Izquierda: La fuente genera una señal triangular (arriba, serie temporal, abajo, contenido espectral). La señal es filtrada por un tubo cerrado-abierto (respuesta en frecuencia en el panel del medio). Dada la respuesta en frecuencia del tubo algunas frecuencias del sonido emitido por la fuente se amplificaran y otras se atenuaran. La salida del filtro se muestra en el panel de la derecha. Se observa un cambio en el contenido espectral (abajo) y en la serie temporal (arriba). Adaptado de (Mindlin and Laje, 2005).

2009).

Dado que el tracto vocal puede pensarse como alteraciones de un tubo uniforme, este se suele modelar como un único tubo ó como una serie de tubos conectados en serie. Si lo consideramos como un único tubo, podemos pensarlo como un tubo con la fuente sonora en un extremo y el otro extremo abierto a la atmósfera. Dentro del tubo existirán dos ondas, una que viaja hacia adelante y otra que viaja hacia atrás. Las mismas corresponden respectivamente a las fluctuaciones de presión introducidas por la fuente y la reflexión de las mismas en el otro extremo. La ecuación que define la dinámica es:

$$a(t) = p_i(t) - \alpha a(t - \tau),$$

donde  $a(t)$  corresponde a la presión al inicio del tubo,  $p_i(t)$  a las fluctuaciones introducidas por la fuente,  $\alpha$  al coeficiente de reflexión entre el tubo y el exterior y  $\tau = \frac{2L}{c}$  al tiempo que tarda la onda en viajar el tubo ida y vuelta; en estas expresiones  $L$  es la longitud del tubo y  $c$  la velocidad del sonido.

La física de este modelo nos indica que el mismo tendrá máximos y mínimos de transferencia para distintas frecuencias. La primera resonancia está dada cuando la longitud del tubo es igual a longitud de onda dividido 4. La misma queda,

$$f_r = \frac{c}{4L} \quad (2.6)$$

Considerando la ubicación de los mínimos se puede aproximar el ancho de banda del filtro como:

$$BW = \frac{c}{8L} \quad (2.7)$$

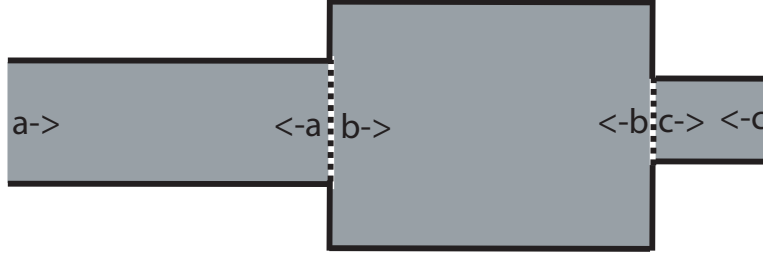


Figura 2.6: Esquema del tracto vocal para un modelo de varios tubos. Cada tubo representa una estructura anatómica del tracto vocal. En cada uno de ellos existe una onda incidente y reflejada producto de transmisiones y reflexiones en las interfases.

El modelo de varios tubos permite incorporar con mayor detalle la morfología del tracto. Este modelo considera los aportes de tres cavidades conectadas en serie: tráquea, cavidad orofaríngea y pico. Los coeficientes de transmisión y reflexión entre los tubos está dado por,

$$\begin{aligned} r_{ij} &= \frac{A_i - A_j}{A_i + A_j} \\ t_{ij} &= 1 + r_{ij} \end{aligned}$$

donde  $r_{ij}$  y  $t_{ij}$  son los coeficientes de transmisión y reflexión del tubo  $i$  al tubo  $j$ .  $A_i$  y  $A_j$  las áreas transversales de los respectivos tubos.

Nuevamente considerando una onda transmitida y otra reflejada en cada tubo, el modelo de tres tubos (a, b y c) queda,

$$\begin{aligned} a_{\rightarrow}(t) &= p_i(t) + a_{\leftarrow}(t - \tau_1) \\ a_{\leftarrow}(t) &= r_{ab} a_{\rightarrow}(t - \tau_1) + t_{ba} b_{\leftarrow}(t - \tau_2) \\ b_{\rightarrow}(t) &= t_{ab} a_{\rightarrow}(t - \tau_1) + r_{ba} b_{\leftarrow}(t - \tau_2) \\ b_{\leftarrow}(t) &= r_{bc} b_{\rightarrow}(t - \tau_2) + t_{cb} c_{\leftarrow}(t - \tau_3) \\ c_{\rightarrow}(t) &= t_{bc} b_{\rightarrow}(t - \tau_2) + r_{cb} c_{\leftarrow}(t - \tau_3) \\ c_{\leftarrow}(t) &= \alpha c_{\rightarrow}(t - \tau_3) \end{aligned}$$

donde  $X_{\rightarrow}$  y  $X_{\leftarrow}$  son las ondas en los extremos de los tubos,  $\tau_i = \frac{L_i}{c}$  el tiempo de viaje para cada tubo y  $\alpha$  el coeficiente de reflexión con la atmósfera. Se puede ver un esquema del tracto en la Figura 2.6.

La principal ventaja de este modelo más completo del tracto vocal es la posibilidad de incorporar variables que detallan la anatomía y dinámica del tracto. Esto permite representar las modificaciones ejercidas por el ave en su tracto durante el canto (por ejemplo, elongaciones de la tráquea o apertura/cierre del pico). De ese

modo, los parámetros del modelo ( $r_{ij}, t_{ij}, \alpha$  y  $\tau_i$ ) dejan de ser valores fijos para transformarse en funciones del tiempo que dependen de las fluctuaciones de sus variables fisiológicas respectivas.

### 2.2.2. Validación experimental de los modelos

Para validar experimentalmente el modelo, Mindlin y colaboradores (Mindlin et al., 2003) registraron las variables fisiológicas asociadas a los parámetros dinámicos del modelo. Utilizando dichas señales fisiológicas como parámetros e integrando el modelo se obtuvo canto sintético comparable con el efectivamente vocalizado por un Cardenal, cuyo canto se caracteriza por ser prácticamente tonal. Las variables fueron registradas mientras el ave cantaba.

Las variables registradas fueron: la presión en los sacos aéreos y la actividad electromiográfica de los músculos  $vS$  y  $dTB$ . Dichas variables corresponden respectivamente a los parámetros:  $p_{sub}$ ,  $k$  y  $g_0$  del modelo representado por 2.4 y 2.5. Para reconstruir la actividad muscular a partir de los registros electromiográficos, dichas señales fueron rectificadas y posteriormente suavizadas. Para transformar los registros de presión y actividad muscular en los parámetros correspondientes se utilizó una transformación lineal. Dicha transformación se ajustó usando los datos de una primera vocalización y luego se mantuvo la transformación para el resto del canto. En la Figura 2.7 se observan sonogramas de las sílabas experimentales y sintéticas. Este resultado refuerza las hipótesis planteadas y reafirma la identificación de variables relevantes para la síntesis de canto.

Gardner y colaboradores realizaron otro tipo de validación experimental para el modelo (Gardner et al., 2005). Una de las preguntas que respondieron es si las aves interpretan este canto sintético como generado por un conoespecífico. Para ello expusieron a canarios juveniles a una grabación de canto sintético acústicamente similar al de canarios adultos pero diseñado con una secuencia de sílabas distinta a la típicamente encontrada en esta especie. Los juveniles lograron copiar dicha canción adoptándola como si fuese originada en un tutor conoespecífico.

La revisión realizada en este capítulo resulta una ilustración de los aspectos importantes aportados por la física y la biología al estudio del canto de las aves. La visión cuantitativa aportada por la física resulta importante para develar los mecanismos involucrados en los fenómenos observados. En el próximo capítulo profundizaremos en esa dirección.



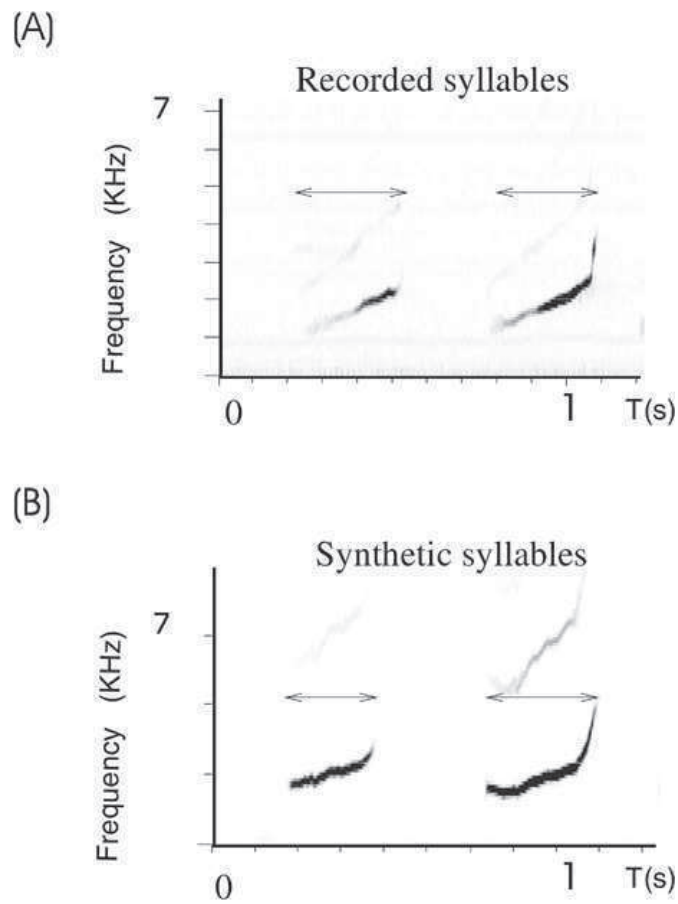


Figura 2.7: (a) Sonograma de un fragmento del canto de cardenal (b) Sonograma de canto sintético generado cuando el modelo es conducido por las variables fisiológicas registradas. Tomado de [Mindlin and Laje \(2005\)](#) con autorización de los autores.

# Diamante Mandarín

## Índice

<b>3.1. Estudio experimental</b> . . . . .	<b>24</b>
<b>3.2. Un nuevo mecanismo dinámico</b> . . . . .	<b>29</b>
<b>3.3. Análisis dinámico</b> . . . . .	<b>35</b>
<b>3.4. Variables de control de las vocalizaciones</b> . . . . .	<b>37</b>

El Diamante Mandarín (*Taeniopgia guttata*) es un ave que pertenece al orden de los Paseriformes, a la familia de los Estríldidos y al género *Poephila*. Es un ave de pequeño tamaño con un peso entre los 10 y 15 g y longitud entre 10 y 13 cm. Habita principalmente en Australia, con preferencia por las zonas desérticas.

El Diamante Mandarín es un pájaro muy sociable que en la naturaleza se reúne en bandadas de decenas de parejas. Una vez que forma una pareja, ésta se mantiene hasta que uno de los dos ejemplares muere. En algunas ocasiones especiales pueden llegar a mantener relaciones polígamicas entre varios pájaros, donde un macho puede cuidar simultáneamente dos puestas con sus respectivas hembras. El proceso de cría de este ave es extremadamente rápido, se mantiene durante todo el año, pudiendo producir múltiples camadas. Por este motivo y su buena adaptabilidad al cautiverio, es la especie más popular en los laboratorios de investigación en canto de aves.

Muchas aves canoras producen sonidos prácticamente tonales, como los canarios (*Serinus canaria*), los chingolos (*Zonotrichia capensis*) y los cardenales (*Cardinalis cardinalis*). Por otro lado, el canto del Diamante Mandarín alterna sonidos tonales de alta frecuencia con vocalizaciones de baja frecuencia con mucho contenido espectral. Los modelos matemáticos presentados en el capítulo anterior reproducen exitosamente canto de característica tonal (Gardner et al., 2001; Laje et al., 2002; Mindlin et al., 2003) pero no logran generar canto de ave con una estructura espectral más compleja.

En este capítulo, nuestra intención es probar la hipótesis que la riqueza espectral del canto del diamante mandarín emerge de la dinámica altamente no-lineal de su aparato fonador, la siringe. Análogamente a lo que sucede en los distintos registros vocales en humanos, el contenido espectral de los sonidos es, no sólo un efecto inducido por el filtrado del tracto superior sino que depende principalmente de los distintos patrones vibratorios de las cuerdas vocales.

### 3.1. Estudio experimental

Los cantos de seis diamante mandarín machos adultos fueron grabados usando un micrófono (Audiotechnica AT 8356). Simultáneamente fue medida la presión en los sacos aéreos mientras el ave cantaba espontáneamente. La medición de la presión se realizó insertando una cánula flexible (Silastic laboratory tubing, 1.65 mm de diámetro exterior) a través de la pared abdominal caudal a la última costilla. De este modo, la cánula ingresa solo unos milímetros dentro del saco aéreo torácico. El extremo libre de la cánula es conectado a un sensor de presión piezoresistivo (Fujikura model FPM-02PG) el cual se coloca en una mochila en la espalda del ave (Goller et al., 2004).

Los cantos de los machos son una secuencia estereotipada de elementos sonoros (sílabas). Estas sílabas varían mucho en su frecuencia fundamental (desde 400 hasta 7000  $Hz$ ; Un canto típico se muestra en la figura 3.1). A diferencia de las vocalizaciones tonales producidas por otras aves, los Diamante Mandarin tienen una gran variabilidad en el contenido espectral de sus sílabas. Su canto exhibe desde sonidos de baja frecuencia con una presencia completa en sus armónicos (harmonic stacks), hasta elementos de alta frecuencia con estructura tonal (como ejemplo ver paneles b y c de la Figura 3.1, correspondientes a las sílabas b y c indicadas en el panel a).

Interesantemente, el contenido espectral está relacionado con la frecuencia fundamental del sonido. Vocalizaciones de baja frecuencia (frecuencia fundamental inferior a 1.5  $KHz$ ) son espectralmente ricas, mientras que sonidos de alta frecuencia fundamental prácticamente no contienen armónicos (ver espectros respectivos en los paneles de la derecha de la figura 3.1)

Debido a la gran variabilidad en frecuencia fundamental y amplitud de las vocalizaciones emitidas por el ave, resulta un desafío comparar entre sí el contenido espectral de distintas vocalizaciones. Es necesario generar una cuantificación de dicho contenido espectral que no dependa de la frecuencia fundamental ni de la amplitud absoluta. Para ello, definiremos un índice de contenido espectral SCI (de sus siglas en inglés, *spectral content index*). Dicho índice se calcula a partir de parámetros extraídos de la transformada rápida de Fourier (FFT) y de las series temporales de cada una de las sílabas. Estos parámetros son: la frecuencia fundamental media (AFF) y la frecuencia espectral media (MSF).

El parámetro AFF se obtuvo del siguiente modo: en el caso de vocalizaciones de baja frecuencia se contabiliza la tasa de pulsos a lo largo de serie temporal de la sílaba y para las vocalizaciones de alta frecuencia se utilizó el primer pico en la FFT. Cuando el registro sonoro del canto era ruidoso, resultó útil realizar la identificación de la frecuencia fundamental a partir de los registros de presión filtrados adecuadamente. Los registros de presión tienen montadas oscilaciones de baja amplitud correspondientes a las oscilaciones de la onda sonora que emite el pájaro. Para recuperar dichas oscilaciones se filtra el registro de presión con un filtro pasa-banda con frecuencias de corte inferior igual a 150  $Hz$  y superior igual a 4000  $Hz$ . En la Figura 3.2 puede observarse este procedimiento.

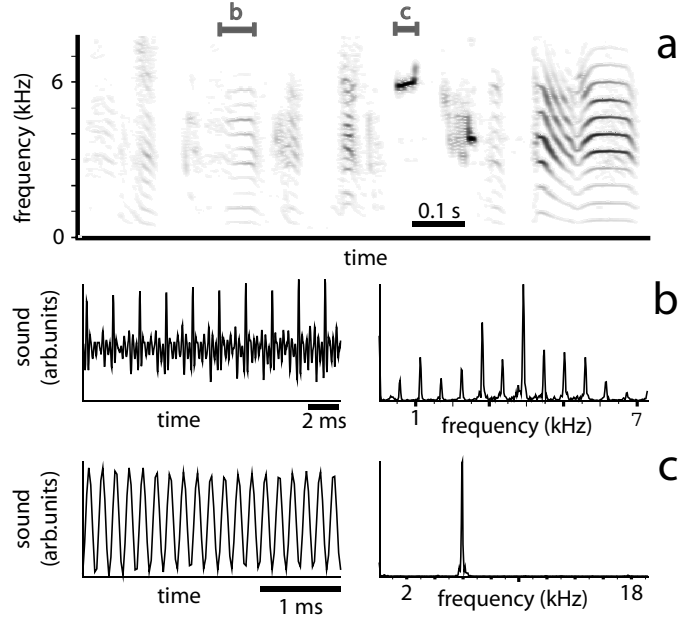


Figura 3.1: Panel superior espectrograma de un canto típico de Diamante Mandarín . En los paneles inferiores se analizan los distintos tipos de sílabas indicados por las barras en el panel superior. (b) Serie temporal y transformada de Fourier de un segmento frecuencia fundamental 560 Hz y alto contenido espectral. (c) Serie temporal y transformada de Fourier de un segmento frecuencia fundamental 5940 Hz y bajo contenido espectral.

El otro parámetro buscado, MSF, se define como la suma de los productos de las frecuencias  $\omega_i$  por la energía correspondiente a dicha frecuencia  $\varepsilon_i$  dividido por la energía total ( $E = \sum_i \varepsilon_i$ ):

$$MSF = \sum_i \frac{\omega_i \varepsilon_i}{E}. \quad (3.1)$$

De modo de cuantificar el contenido espectral de cada sílaba independientemente de su frecuencia fundamental definimos el índice de contenido espectral (SCI) como el cociente entre MSF y AFF.

$$SCI = \frac{MSF}{AFF}. \quad (3.2)$$

Este coeficiente nos permite comparar sílabas muy variadas en frecuencia fundamental. El cociente resulta igual a 1 cuando el sonido es tonal puro y progresivamente mayor a 1 a medida que el sonido tiene más energía en sus armónicos superiores.

En la Figura 3.3 se muestra el cálculo del contenido espectral en función de la frecuencia fundamental para 6 individuos generando distintos tipos de sílabas. En primera instancia puede afirmarse que sílabas de baja frecuencia fundamental tienen alto contenido espectral, mientras que sílabas de alta frecuencia fundamental son prácticamente tonales ( $SCI \sim 1$ ). Esto se corresponde con lo originalmente

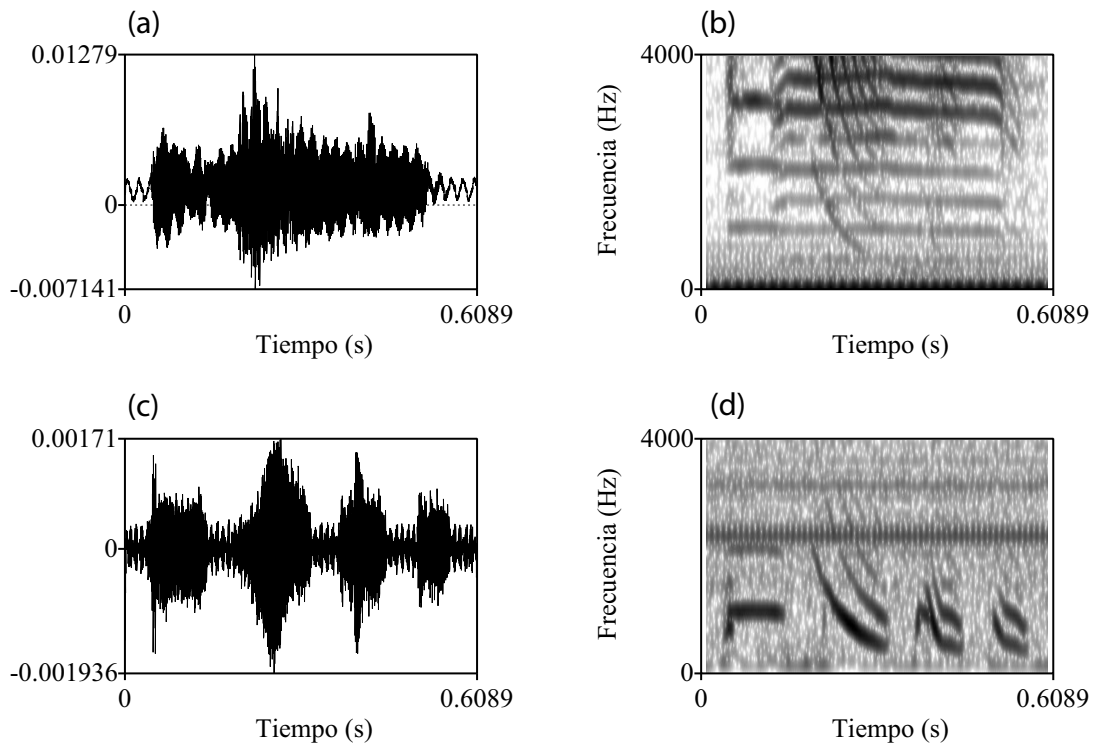


Figura 3.2: Identificación de la frecuencia fundamental a partir del registro de presión. Se muestra el procedimiento en una grabación donde la vocalización del ave coincide con la de otra ave de fondo, este ejemplo es importante ya que descarta el crosstalk entre canales (a) Registro sonoro de las vocalizaciones (b) Sonograma del registro sonoro, en el mismo se observan las vocalizaciones simultaneas de dos aves (c) Registro temporal de la señal de presión luego de pasarle el filtro pasa banda (d) Sonograma del registro de presión filtrado. En mismo desaparece el aporte de la segunda ave y se puede identificar más sencillamente la frecuencia fundamental de las sílabas.

observado en los sonogramas donde las vocalizaciones de baja frecuencia presentan mayor energía distribuida en sus armónicos.

Por otro lado, es muy importante indicar que aun siendo extraídos de cantos de aves distintas, los puntos se organizan a largo de una curva unidimensional en el espacio (*AFF, SCI*) sugiriendo una relación funcional sencilla válida a lo largo de un espacio amplio de frecuencias. La variabilidad de contenido armónico en las vocalizaciones y el hecho que las sílabas de baja frecuencia sean generadas mediante una dinámica labial pulsátil (Jensen et al., 2007; Goller and Cooper, 2004) sugiere fuertemente que la relación sistemáticamente observada entre frecuencia fundamental y tonalidad emerge del mecanismo responsable de la generación de sonido.

Para probar esta hipótesis, realizamos una exploración teórica de la dinámica labial de modo de identificar un mecanismo dinámico capaz de ser responsable de la relación observada entre contenido espectral y frecuencia fundamental. Posteriormente, nuestro objetivo es verificar si dicho mecanismo dinámico se encuentra presente en un modelo de baja dimensión de la dinámica labial.

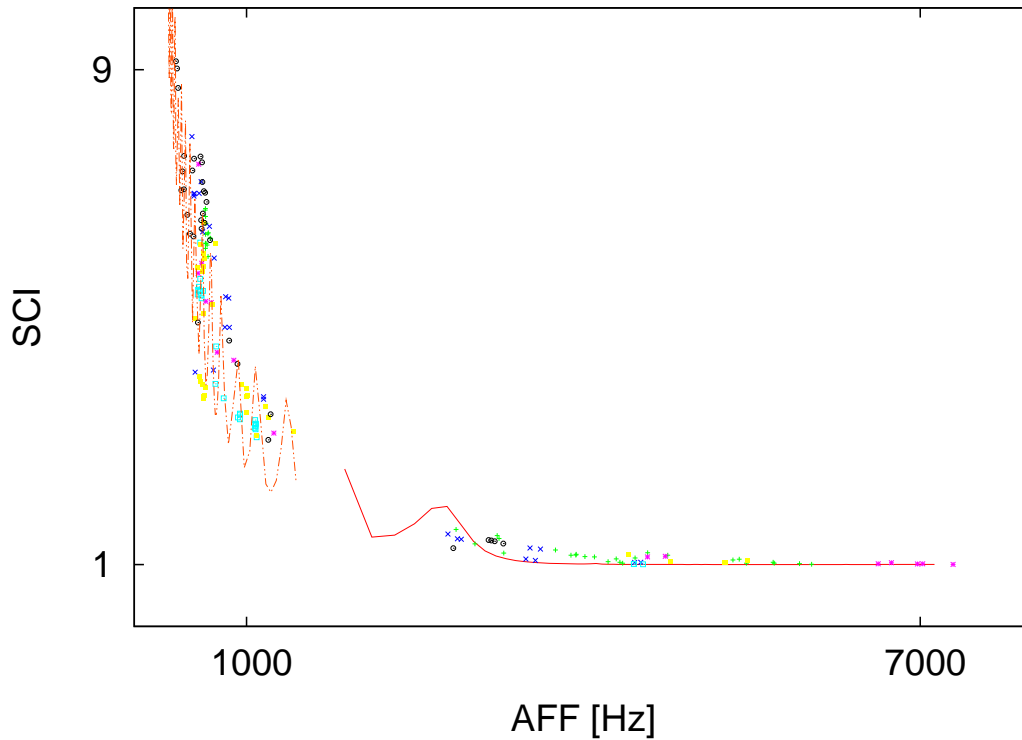


Figura 3.3: El contenido espectral de las sílabas y su frecuencia fundamental están correlacionados. Se muestra el análisis de 172 sílabas realizadas por 6 Diamante Mandarín. Para cada sílaba se calcula el índice de contenido espectral (SCI) y la frecuencia fundamental media (AFF). Cada sílaba analizada es representada por un punto, distintos tipos de punto corresponden a distintas aves. La línea punteada se obtiene integrando el modelo para distintos valores del parámetro  $p_{sub}$  (presión en los sacos aéreos) de modo de generar sonidos de frecuencia fundamental menor a  $1.5\text{ KHz}$ . Luego dichos sonidos son filtrados con un tubo de longitud  $20\text{ mm}$  y un coeficiente de reflexión  $\alpha=0.95$ . La línea continua se genera integrando el modelo modificando los parámetros  $p_{sub}$  y  $g_0$  (actividad muscular) y luego filtrando usando el mismo tubo.

## 3.2. Un nuevo mecanismo dinámico

Como se indicó en el capítulo 2, tanto en las aves como en humanos, la generación de sonido requiere la existencia de oscilaciones (Mindlin and Laje, 2005). La teoría de sistemas dinámicos permite la clasificación de modos mediante los cuales un sistema físico atraviesa un cambio cualitativo en su comportamiento. Dichos cambios se producen en función de la modificación de sus parámetros (Mindlin et al., 1991; Strogatz, 2001). En nuestro caso nos interesa, los tipos de transición que llevan, a los labios siringeos, de un estado estacionario estable hacia uno oscilatorio.

En el modelo presentado en el capítulo anterior las oscilaciones se encienden mediante una bifurcación de Hopf. Como mencionamos oportunamente, las oscilaciones nacidas en dicha bifurcación se caracterizan por ser tonales y por aparecer con una frecuencia bien definida.

Otra de las maneras mediante las que el sistema puede transicionar de un estado estacionario a uno oscilatorio es mediante la bifurcación conocida como nodo-silla en ciclo limite (SNILC, por sus siglas en ingles) (Strogatz, 2001). Esta bifurcación ha sido caracterizada también en otros sistemas biológicos, por ejemplo en la excitabilidad tipo 1 de las neuronas (Izhikevich, 2006).

En esta bifurcación, en el estado estacionario existen 3 puntos fijos: un atractor, un silla y un repulsor como se muestra en la Figura 3.4(a). El sistema tiene la particularidad que la variedad estable del atractor se conecta con la variedad inestable del repulsor, formando así un círculo invariante (“lo que pasa en el círculo queda en el círculo”). La bifurcación surge de la colisión del punto silla y el atractor, generando así un ciclo limite, esto esta representado respectivamente por los paneles (b) y (c) en la Figura 3.4. Las oscilaciones nacen con frecuencia cero y amplitud finita. Estas características estan dadas debido a que justo después de la bifurcación el sistema oscila pasando un largo periodo de tiempo en la región del espacio de fases  $(x, dx/dt)$  donde la aniquilación sucedió. Esto se debe a que en ese lugar existe un remanente del atractor o “fantasma” de la bifurcación (Figura 3.4(c)). Bajo estas condiciones el sistema genera oscilaciones de gran contenido espectral (ver Figura 3.6( $B_1$ )). A medida que el sistema se aleja de la bifurcación el efecto del “fantasma” es menor y las oscilaciones son cada vez más tonales (Figura 3.6( $B_2$ )). Por ende, este mecanismo es un buen candidato para replicar las características espectrales observadas. Más aun, el mismo puede ser identificado en un modelo físico de la dinámica labial.

Este modelo ha sido ya presentando en el capítulo anterior y fue usado con éxito para sintetizar sonido tanto para humanos como para aves. Se realizo una modificación en el modelo de modo que cumpla con los requisitos dinámicos para tener una bifurcación SNILC. Para ello, incorporando términos no-lineales, haremos una expansión en la descripción de las fuerzas elásticas y de disipación. De este modo, la ecuación dinámica para la variable  $x$ , que describe la posición media de los labios respecto a su posición de equilibrio queda:



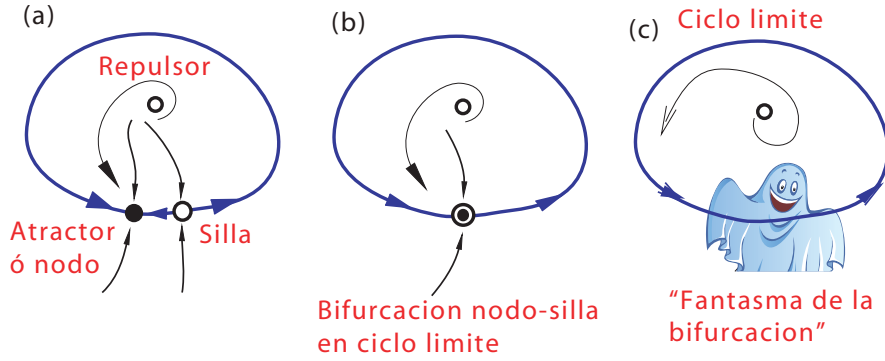


Figura 3.4: Bifurcación nodo-silla en ciclo limite. (a) estado estacionario, existen 3 puntos fijos: repulsor, atractor y silla, estos dos últimos forman un ciclo invariante. (b) la bifurcación sucede cuando chocan el punto silla y el atractor, formando así un ciclo limite. (c) cerca de la bifurcación existe un remanente o “fantasma” de la bifurcación, el cual enlentece el recorrido en dicho segmento de la trayectoria en el ciclo límite.

$$\frac{dx}{dt} = y \quad (3.3)$$

$$m \frac{dy}{dt} = -k(x)x - \beta(y)y - cx^2y - g_0 + a_{lab}p_{sub} \frac{\Delta a + 2\tau y}{a_{01} + x + \tau y} \quad (3.4)$$

En el resto de esta tesis llamaremos a este modelo, “Modelo Físico” de la siringe. La primera modificación introducida corresponde a la fuerza de restitución, se reemplaza por el parámetro  $k$  por una función no-lineal  $k(x) = k_1 + k_2x^2$ . La segunda modificación es en la disipación, reemplazando el parámetro  $b$  con la función  $\beta(y) = \beta_1 + \beta_2y^2$ .

Mediante la integración numérica de estas ecuaciones modelamos las fluctuaciones de presión ( $P_i$ ), introducidas por la fuente, en la base de tracto vocal. Esas ondas sonoras son filtradas por el tracto, el cual simplificamos como un tubo de longitud  $L$  y coeficiente de reflexión  $\alpha$  (ver esquema de la siringe en la Figura 3.5). La fluctuaciones de presión en la base de la traquea ( $a(t)$ ) a tiempo  $t$  están dadas por:

$$a(t) = P_i(t) - \alpha a(t - \frac{2L}{v}), \quad (3.5)$$

donde  $\alpha$  representa el coeficiente de reflexión al final de la traquea y  $v$  es la velocidad del sonido.

De este modo, modelamos el sonido en términos de la fuente, cuya dinámica esta determinada por las ecuaciones 3.3 y 3.4, y un filtro que posteriormente resaltará algunas frecuencias y atenuará otras.

Implementamos matemáticamente este modelo y buscamos una región en el espacio de parámetros donde la fuente sonora presente el comportamiento dinámico

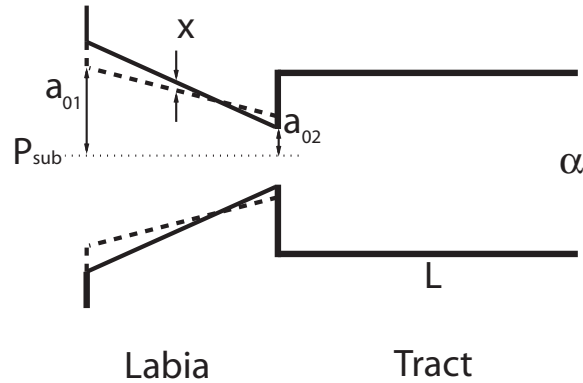


Figura 3.5: Esquema del modelo siringe-filtro.  $p_{sub}$  corresponde a la presión sub-siringea,  $x$  el punto medio de desplazamiento de los labios,  $a_{01}$  y  $a_{02}$  la separación media en el estado de reposo,  $L$  la longitud del tubo y  $\alpha$  el coeficiente de reflexión al final de la tráquea.

previamente descrito, por ejemplo, el nacimiento de oscilaciones con baja frecuencia y alto contenido espectral.

En la Figura 3.6 se muestran simulaciones para tres valores de presión sub-siringea (Puntos A,  $B_1$  y  $B_2$ ). El sistema exhibe una bifurcación SNILC. Para un valor de presión bajo (punto A, de la figura) el sistema no oscila. Tal como se describió anteriormente, en dicha situación un atractor coexiste con un punto silla y un repulsor. La variedad inestable del punto silla se conecta con la variedad estable del atractor. De ese modo, cuando el punto silla y el atractor se aniquilan en el punto de bifurcación se crea un ciclo límite. Los casos  $B_1$  y  $B_2$  muestran como la frecuencia aumenta y el contenido espectral disminuye a medida que  $p_s$  se aleja del punto de bifurcación.

Generamos sonido sintético integrando el modelo dados los siguientes parámetros:  $a_{lab} = 2 \times 10^{-4} \text{ cm}^2$ ,  $m = 0.4 \text{ ng}$ ,  $k_1 = 0.36 \text{ dyn/cm}$ ,  $\beta_1 = 4.44 \times 10^{-5} \text{ dyn s/cm}$ ,  $c = 1.6 \times 10^{-2} \text{ dyn s/cm}^3$ ,  $g_0 = -0.0399 \text{ dyn}$ ,  $\tau = 5 \times 10^{-6} \text{ s}$ ,  $a_{01} = 0.1 \text{ cm}$ ,  $a_{02} = 0.11 \text{ cm}$ ,  $k_2 = 400 \text{ dyn/cm}^3$  y  $\beta_2 = 4 \times 10^{-11} \text{ dyn s}^3/\text{cm}^3$ . El contenido espectral fue calculado para 41 sílabas sintéticas de 114 ms de duración (fueron removidos transitorios iniciales de 114 ms). La integración se realizó con un paso de integración de  $dt = 3.5 \times 10^{-3} \text{ ms}$ . Contenido espectral y frecuencia fundamental fueron analizados del mismo modo que en el caso de los sonidos experimentales. La curva teórica (Línea punteada en la Figura 3.3) ajusta satisfactoriamente los sonidos experimentales de frecuencia fundamental inferior a 1.5 kHz.

Esta curva fue generada sintetizando sonidos con valores de  $p_{sub}$  correspondientes a oscilaciones desde los 30 a hasta los 1500 Hz. A medida que  $p_s$  se aleja del punto de bifurcación, la frecuencia de oscilación aumenta. Para cada uno de estos sonidos, el índice de contenido espectral (SCI) fue calculado a partir de los parámetros MSF y AFF del mismo modo que en el caso experimental.

Los parámetros del tubo fueron buscados de modo que minimicen la distancia cuadrática ( $\chi^2$ ) entre la curva SCI vs AFF experimental y su curva par teórica.

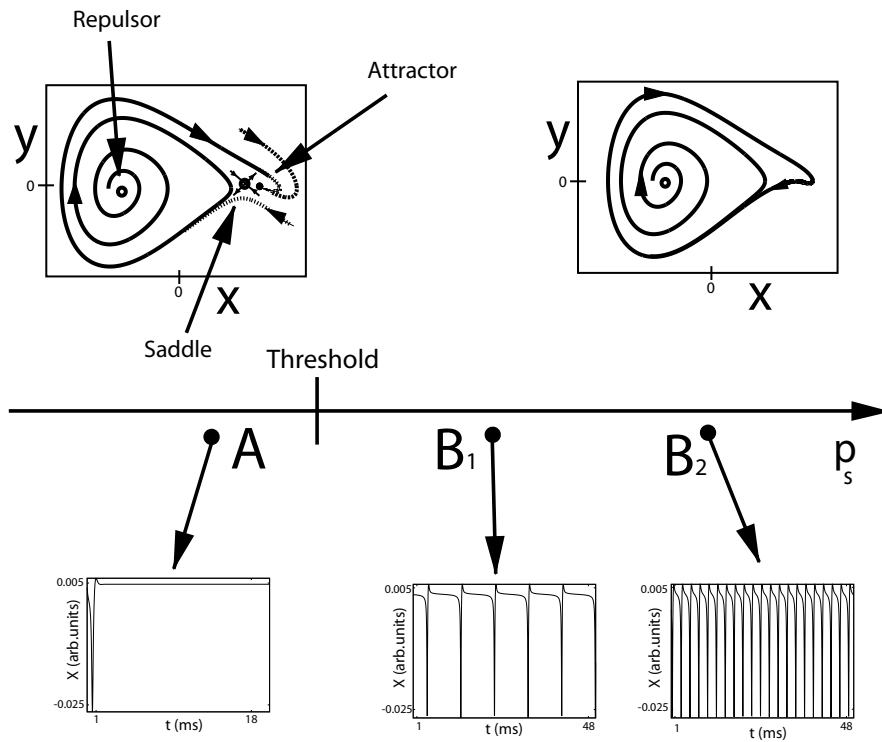


Figura 3.6: El modelo físico puede exhibir una bifurcación de nodo-silla en ciclo límite (SNILC). Esquematisamos la dinámica de la posición media de los labios en función del espacio del parámetro  $p_{sub}$ , el cual está dividido en 2 regiones por una línea de bifurcación. Para valores de  $p_{sub}$  por debajo de la bifurcación existen tres puntos fijos donde el estado estacionario es estable y no ocurren oscilaciones (ver  $A$ ). A medida que  $p_s$  aumenta y cruza la línea de bifurcación nace una oscilación de frecuencia cero con amplitud finita y una traza temporal de tipo pulsátil ( $B_1$ ). Si la presión continúa creciendo y el sistema se aleja aun más de la línea de bifurcación, la frecuencia de oscilación aumenta y la traza temporal se hace más tonal ( $B_2$ ).

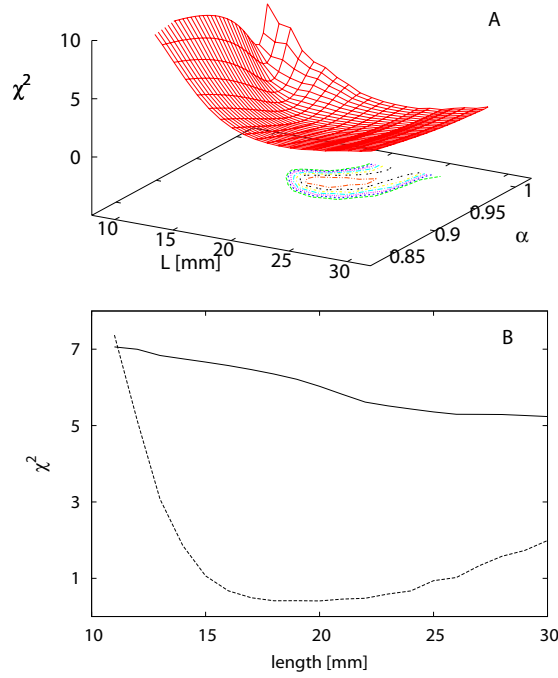


Figura 3.7: (a) Optimización de la distancia entre los valores experimentales y sus pares teóricos ( $\chi^2$ ) en función de los parámetros del tubo  $\alpha$  y  $L$ . El contorno de las líneas de nivel es dibujado en la base de la superficie. (b) En el panel inferior mostramos la relación de  $\chi^2$  y  $L$  dado un  $\alpha$  fijo de 0.95 para los modelos *Hopf* (línea sólida) y *SNILC* (línea punteada).

Los valores obtenidos son:  $\alpha=0.95$  y  $L=20$  mm. En la Figura 3.7(a) se muestra la relación entre  $\chi^2$  y los parámetros del tubo  $\alpha$  y  $L$ .

Los mecanismos dinámicos presentes en los modelos presentados anteriormente en la literatura no nos hubiesen permitido recuperar el comportamiento del índice *SCI* a lo largo de espacio de frecuencias estudiadas. Ello se verificó mediante una hipótesis alternativa: que las oscilaciones nacen en una bifurcación *Hopf* en vez de una *SNILC*. Para verificar esta hipótesis usamos un modelo minimal de la siringe (Laje et al., 2002),

$$\frac{dx}{dt} = y \quad (3.6)$$

$$\frac{dy}{dt} = -k_h x - (p_{sub} - \beta_h)y - c_h x^2 y, \quad (3.7)$$

donde la frecuencia es controlada por el parámetro  $k_h$  y la bifurcación *Hopf* es cruzada cuando  $p_{sub} > \beta_h$ . Generamos sonidos de frecuencias fundamentales desde 350 hasta 7100 Hz barriendo  $k_h$  desde  $5 \times 10^6$  hasta  $2 \times 10^9$ . Los parámetros  $p_{sub}$  y  $\beta_h$  fueron dejados fijos en 8000 y 1000 respectivamente, luego buscamos el mejor filtro para minimizar  $\chi^2$ . Se encontró muy poca dependencia de  $\chi^2$  con los paráme-

tros del filtro y no se pudo recuperar el comportamiento del índice SCI para bajas frecuencias. En la Figura 3.7(b) se muestra la relación entre  $\chi^2$  y  $L$  dados un  $\alpha$  fijo para los modelos *Hopf* y *SNILC*.

De este modo mostramos que el mecanismo SNILC de inicio de oscilaciones permite recuperar la relación entre el contenido espectral y la frecuencia fundamental, por lo menos para sonidos de baja frecuencia. En este mecanismo de producción sonora la presión puede controlar por sí sola la frecuencia fundamental de los sonidos emitidos. Experimentos, en Diamante Mandarin, en los que se eliminó el control motor denervando la siringe muestran una correlación entre frecuencia fundamental y presión (Daley and Goller, 2004). También existe evidencia experimental que este mecanismo de control de la frecuencia por presión esta presente en aves no-oscinas (Amador et al., 2008).

Los sonidos de alta frecuencia podrían ser sintetizados aumentando aun más la presión, pero para alcanzar los valores de frecuencia fundamental experimentales, la presión debería alcanzar valores cuya magnitud se torna irrealista. De hecho, veremos en la sección 3.4 que sílabas de baja y alta frecuencia son vocalizadas con valores similares de presión sub-siringea.

De modo de sintetizar sonidos de alta frecuencia, barrimos los parámetros  $g_0$  y  $p_{sub}$  desde  $-0.04 \text{ dyn}$  a  $-0.0328 \text{ dyn}$  y  $2004 \text{ Pa}$  a  $2010 \text{ Pa}$  respectivamente, correspondiendo a oscilaciones desde  $1.5 \text{ kHz}$  a  $6.3 \text{ kHz}$ . Este enfoque es consistente con las observaciones experimentales, en las que no se observa actividad ventral en los sonidos de baja frecuencia fundamental (Vicario, 1991). Los datos fueron analizados como en el caso de los de baja frecuencia. Se sintetizaron 100 sílabas de  $16,4 \text{ ms}$  de duración (descartando un transitorio inicial de  $16,4 \text{ ms}$ . El paso de integración fue  $dt = 2 \times 10^{-3} \text{ ms}$ ). El SCI se calculó para cada una en función de la frecuencia fundamental y se encuentra graficado con una línea continua en la Figura 3.3. La curva teórica se ajusta satisfactoriamente con los sonidos experimentales de alta frecuencia.

### 3.3. Análisis dinámico

Amador y Mindlin realizaron un análisis dinámico del modelo propuesto (Amador and Mindlin, 2008). El análisis se realizó en términos de los parámetros que fueron encontrados relevantes en trabajos previos para la producción de canto en aves oscinas (Gardner et al., 2001; Mindlin et al., 2003). En la Figura 3.8 se muestra una región en el espacio de parámetros  $(p_s, k_1)$ . Dicha región puede dividirse en dos, separando el espacio donde existen tres puntos fijos (región sombreada) de donde existe un único punto fijo. La frontera entre estas dos regiones corresponde a una bifurcación nodo-silla. La misma esta representada por dos curvas con forma de cúspide (líneas punteadas verdes). También se identifica la ubicación de una bifurcación de Hopf, la misma representa un mecanismo importante de encendido de oscilaciones. En la Figura se representa como una línea sólida roja. Ambas curvas de bifurcaciones (Hopf y Nodo-silla) coinciden tangencialmente en un punto con una tercera bifurcación: la bifurcación Homoclínica. En esta bifurcación el ciclo límite choca con el punto silla aniquilando así las oscilaciones. La Homoclínica se indica en la Figura con una línea rayada azul. Este punto de coincidencia de las tres bifurcaciones es una bifurcación de co-dimensión 2 (es necesario ajustar dos parámetros para que la bifurcación suceda). Dicha bifurcación es conocida como bifurcación de Takens-Bogdanov (Guckenheimer and Holmes, 1990).

A partir de las curvas de estas bifurcaciones pueden definirse en el espacio de parámetros regiones con distintos comportamientos dinámicos. En la Figura 3.8 se enumeran dichas regiones y se presentan integraciones numéricas del modelo con valores de los parámetros correspondientes a cada una de ellas.

En las regiones 1 y 2 existe un único punto fijo. Ambas regiones están separadas por la bifurcación de Hopf. La estabilidad de dicho punto fijo cambia de estable (región 1) a inestable (región 2). En las regiones 3-7 (área sombreada) existen tres puntos fijos. Las mismas están separadas de las regiones 1-2 por la bifurcación nodo-silla. En la zona 3 existen tres puntos fijos: dos estables y uno silla. En esta zona existe una bi-estabilidad en el espacio de fases (dependiendo de las condiciones iniciales el sistema se estabilizara en uno de los dos puntos fijos estables). La misma se encuentra separada de la zona 4 por la bifurcación de Hopf, la cual cambia la estabilidad de uno de los puntos fijos estables generando un ciclo límite estable alrededor del mismo. La zona 4 se separa de la zona 5 mediante la bifurcación homoclínica. En dicha bifurcación el ciclo limite se destruye al tocar el punto silla.

La zona 5 es de especial interés, ya que la variedad inestable del punto silla se contacta con la variedad estable del atractor. Esto, como ya mencionamos anteriormente, forma un círculo invariante. Cuando colisionan el punto silla con el atractor (bifurcación nodo-silla, pasaje de zona 5 a zona 2) nacen oscilaciones en un ciclo límite con las características ya mencionadas de la bifurcación SNILC. La zona 6 es separada de la 5 por una nueva bifurcación homoclínica y la 6 de la 7 por otra bifurcación de Hopf.

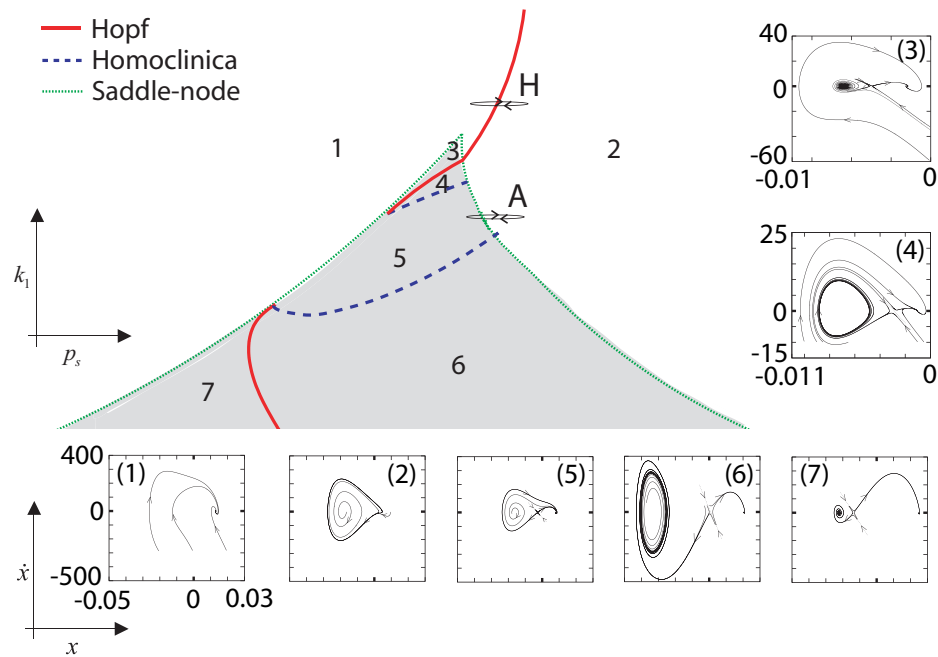


Figura 3.8: Diagrama de bifurcaciones del modelo físico en el espacio de parámetros ( $p_s$ ,  $k_1$ ). El diagrama de bifurcaciones corresponde  $p_s$  entre 1852 y 2084 Pa y  $k_1$  entre 0,16 y 0,52  $dyn/cm^3$ . En esta región de parámetros ocurre una bifurcación de Takens-Bogdanov, donde una bifurcación nodo-silla (línea punteada) toca tangencialmente una bifurcación de Hopf (línea continua) y una bifurcación homoclínica (línea rayada). Dentro de la región sombreada existen tres puntos fijos y fuera de ella solo uno. El camino indicado por A pasa por un bifurcación *SNILC* y el indicado por H por una Hopf. La integración del modelo con parámetros correspondientes a cada región numerada se muestra en los correspondientes recuadros. Tomado de [Amador and Mindlin \(2008\)](#) con la autorización de los autores

### 3.4. Variables de control de las vocalizaciones

Es importante poder identificar las variables fisiológicas usadas por el ave para poder controlar la frecuencia fundamental y el contenido espectral. Nuestra hipótesis es que en el caso de las sílabas de baja frecuencia, la frecuencia fundamental es controlada fundamentalmente por la presión, lo cual es consistente con reportes en la literatura que describen ausencia de actividad muscular siríngica para dichas sílabas (Vicario, 1991). En el caso de las sílabas de alta frecuencia, es necesaria la presencia de actividad muscular de modo de alcanzar dichos registros. Veremos un poco más al respecto en el próximo capítulo. Por último, en el caso de la relación entre la frecuencia fundamental y el contenido espectral, nuestra hipótesis es que dicha relación es fundamentalmente consecuencia directa del mecanismo dinámico responsable de la fonación y no existiría un control activo de parte del ave.

Para generar confianza en estas hipótesis, en la Figura 3.9(a) analizamos la relación entre la frecuencia fundamental del sonido y la presión en los sacos aéreos. Puntos diferentes corresponden a sílabas distintas. Observar que en el rango de baja frecuencia existen cuatro sílabas distintas (las cuales se encuentran dentro del rectángulo punteado de la Figura 3.9(a)). Los sonidos de alta frecuencia pueden ser generados con el mismo rango de presión que los presentados en algunas de las sílabas de baja frecuencia. Por ende, la presión en los sacos aéreos no puede determinar por sí sola las melodías de alta frecuencia: es necesaria la presencia de otro mecanismo de control.

En el modelo, los sonidos de alta frecuencia pueden ser sintetizados cambiando los parámetros correspondientes a la actividad de los músculos siríngicos. De modo de revelar en este rango de frecuencias la relación entre frecuencia fundamental y contenido espectral, se realizaron simulaciones en las cuales se cambiaron linealmente los parámetros que representan la presión en los sacos aéreos y la actividad muscular dorsal. Dichas simulaciones generaron vocalizaciones de alta frecuencia, las cuales fueron analizadas del mismo modo que los sonidos experimentales.

De ese modo, se lograron sonidos de alta frecuencia con los mismos niveles de presión que en las vocalizaciones de baja frecuencia. En la Figura 3.9(b) se muestra cómo, utilizando el modelo físico, es posible replicar todos los sonidos. En el caso de las vocalizaciones de baja frecuencia, se replican variando únicamente  $p_{sub}$ . Por otro lado, en el caso de las vocalizaciones de alta frecuencia fundamental, se logra variando simultáneamente  $p_{sub}$  y  $g_0$ . Los caminos recorridos en el espacio de parámetros  $(p_{sub}, g_0)$  son los siguientes: cuadrados (2000.8 Pa, -0.0384 dyn) a (2001.4 Pa, -0.0384 dyn); triángulos inferiores (2001.6 Pa, -0.0356 dyn) a (2001.8 Pa, -0.0356 dyn); círculos (2001.2 Pa, -0.0399 dyn) a (2001.4 Pa, -0.0359 dyn); triángulos superiores (2001.2 Pa, -0.0396 dyn) a (2001.3 Pa, -0.0359 dyn).

Los sonidos sintéticos de baja frecuencia fundamental fueron generados, y posteriormente filtrados, modificando el parámetro  $p_{sub}$  cerca del lugar de la bifurcación *SNILC*. Dichos sonidos son presentados en la Figura 3.3, donde se muestra un ajuste teórico a los sonidos experimentales de baja frecuencia. Similarmente, modificando el parámetro  $g_0$  y filtrando con el mismo filtro que en el caso anterior, se



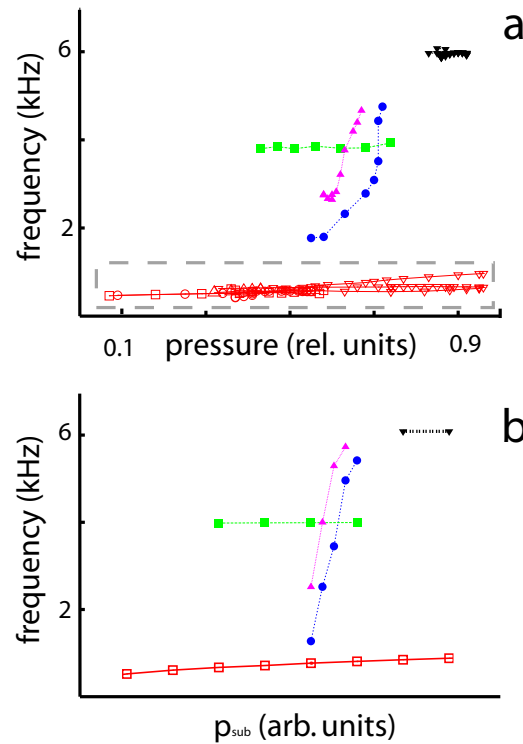


Figura 3.9: Sonidos experimentales y sintéticos de un ave. La medición simultánea de sonido y presión en los sacos aéreos ( $p_s$ ) mientras el ave canta espontáneamente, permite analizar la relación entre la frecuencia fundamental del sonido y la presión en los sacos aéreos. Sílabas diferentes son representadas con puntos distintos. Se puede observar que en el rango de baja frecuencia (rectángulo punteado en a) existen cuatro sílabas distintas. Dichas sílabas, pueden ser replicadas con el modelo físico variando únicamente  $p_s$ . El resto de las sílabas (cuya frecuencia fundamental es mayor a  $1.5 \text{ kHz}$ ), pueden ser también sintetizadas controlando en forma simultánea  $p_s$  y  $g_0$  o  $k$  (quienes representan la actividad de los músculos siringeos). (b) Los caminos en el espacio de parámetros ( $p_s$ ,  $g_0$ ) para generar las sílabas sintéticas con frecuencia fundamental alta son: para los cuadrados de ( $2000.8 \text{ Pa}$ ,  $-0.0384 \text{ dyn}$ ) a ( $2001.4 \text{ Pa}$ ,  $-0.0384 \text{ dyn}$ ); para los triángulos inferiores de ( $2001.6 \text{ Pa}$ ,  $-0.0356 \text{ dyn}$ ) a ( $2001.8 \text{ Pa}$ ,  $-0.0356 \text{ dyn}$ ); para los círculos de ( $2001.2 \text{ Pa}$ ,  $-0.0399 \text{ dyn}$ ) a ( $2001.4 \text{ Pa}$ ,  $-0.0359 \text{ dyn}$ ); para los triángulos superiores de ( $2001.2 \text{ Pa}$ ,  $-0.0396 \text{ dyn}$ ) a ( $2001.3 \text{ Pa}$ ,  $-0.0359 \text{ dyn}$ ).

---

generaron sonidos sintéticos de alta frecuencia fundamental, que ajustan exitosamente a los datos experimentales.

En este capítulo se analizó la relación entre el contenido espectral y la frecuencia fundamental del repertorio de vocalizaciones encontradas en el canto del Diamante Mandarín. Análogamente a lo que sucede en humanos, dicha relación depende principalmente del mecanismo de fonación de la fuente (Titze, 1988). Modificando un modelo previamente usado en la literatura, se generaron sonidos que se ajustan con éxito a los vocalizados por el ave. La incorporación de términos no-lineales y la exploración detallada del modelo, permitió reportar un régimen dinámico de soluciones pulsátiles similares a las encontradas en las vocalizaciones de baja frecuencia en los Diamantes Mandarín. De este modo, la fuerte dependencia entre el contenido espectral y la frecuencia fundamental emerge de las propiedades intrínsecas del aparato vocal y no depende de un control neuronal complejo.



# Biomimética vocal

---

## Índice

<b>4.1. Modelo simplificado</b> . . . . .	<b>41</b>
4.1.1. Análisis dinámico . . . . .	42
4.1.2. Escala temporal . . . . .	44
<b>4.2. Siringe analógica</b> . . . . .	<b>49</b>
<b>4.3. Siringe digital</b> . . . . .	<b>54</b>
4.3.1. ¿Qué es un DSP? . . . . .	54
<b>4.4. Desarrollo de canto sintético</b> . . . . .	<b>56</b>
<b>4.5. Instrucciones fisiológicas conduciendo el DSP</b> . . . . .	<b>62</b>

---

En la introducción de esta tesis, clasificamos la biomimética en las corrientes de biorobótica y neuroprostética. Esta última corriente tiene como objetivo la sustitución de una función biológica. En este capítulo, desarrollaremos la implementación de un equipamiento biomimético del aparato fonador del Diamante Mandarín usando el modelo físico desarrollado en el capítulo anterior.

Uno de los objetivos del Laboratorio de Sistemas Dinámicos es identificar las propiedades acústicas que el ave necesita escuchar de sus propias vocalizaciones durante el período sensorimotor. Estos requisitos se podrán estudiar; realizando alteraciones en la estructura acústica del canto propio escuchado por el ave. A estos experimentos se los llama experimentos de feedback auditivo alterado. En esa dirección, se desarrolló el dispositivo presentado en esta tesis. Su flexibilidad y la integración en el mismo del modelo físico de la siringe, lo hace una herramienta ideal para el desarrollo futuro de los experimentos de feedback alterado.

## 4.1. Modelo simplificado

El modelo físico presentado en el capítulo anterior es capaz de reproducir las características acústicas observadas en el canto del Diamante Mandarín (Sitt et al., 2008). Sin embargo, su implementación en un sistema de integración en tiempo real es desafiante, debido a que el mismo es computacionalmente demandante.

Asumimos que las características acústicas que son importantes de preservar (alto contenido espectral para sonidos de baja frecuencia y bajo contenido espectral para sonidos de alta frecuencia), surgen como firmas de las bifurcaciones que definen el comportamiento oscilatorio del modelo. Resulta entonces tentador,

buscar un conjunto de ecuaciones que presenten un diagrama de bifurcaciones similar, es decir, un comportamiento dinámico equivalente, pero cuya integración sea computacionalmente menos demandante.

Para ello, transformaremos las ecuaciones del modelo físico a un conjunto de ecuaciones más sencillas, utilizando la teoría de formas normales (Guckenheimer and Holmes, 1990; Solari et al., 1996). Dicha técnica consiste en eliminar los términos no lineales, no resonantes en el vecindario de una singularidad lineal (en nuestro caso el cruce de las bifurcaciones Hopf, Nodo-silla y Homoclínica). Alrededor de ese punto en el espacio de parámetros, se efectuó una expansión en serie de Taylor de las ecuaciones del modelo físico. Luego, se buscó un cambio de coordenadas no-lineal que eliminase los términos no resonantes.

Las ecuaciones para la forma normal encontrada son:

$$\frac{dx}{d\tau} = z \quad (4.1)$$

$$\frac{dz}{d\tau} = \alpha + \beta x + x^2 - x^3 - xz - x^2 z, \quad (4.2)$$

Donde la variable  $x$  es análoga a su par en el modelo físico, representando el desplazamiento de labios. Los parámetros  $\alpha$  y  $\beta$  son respectivamente función de la presión y de la actividad de los músculos siríngicos ventrales. Dichas funciones quedan definidas por las transformaciones anteriormente mencionadas, que llevan del sistema físico a su forma normal.

En las diferentes regiones del espacio de parámetros  $(\alpha, \beta)$ , el sistema exhibe dinámicas cualitativamente distintas. El sistema dinámico fue diseñado de modo que presente las mismas series de soluciones que las presentadas por el modelo físico. Para ello, se buscó que estén presentes las mismas bifurcaciones (dos nodo-silla formando un cusp con una Hopf y una homoclínica chocando tangencialmente con una de las nodo-silla.). El diagrama de bifurcaciones se muestra en la Figura 4.1, el mismo debe ser comparado con el del sistema físico en la Figura 3.8.

#### 4.1.1. Análisis dinámico

Para realizar el análisis dinámico del sistema propuesto, primero identificaremos la ubicación de los puntos fijos. Igualando las derivadas de las ecuaciones 4.1 y 4.2 a cero, obtenemos,

$$z = 0 \quad (4.3)$$

$$\alpha + \beta x + x^2 - x^3 = 0, \quad (4.4)$$

La cantidad de raíces de la ecuación 4.4 determinará el número de puntos fijos que posee el sistema. En nuestro caso, nos interesa la transición entre la existencia de uno a tres puntos fijos. Dicha transición define la bifurcación de nodo-silla. Esta situación está representada por la condición crítica de dos raíces reales de la ecuación

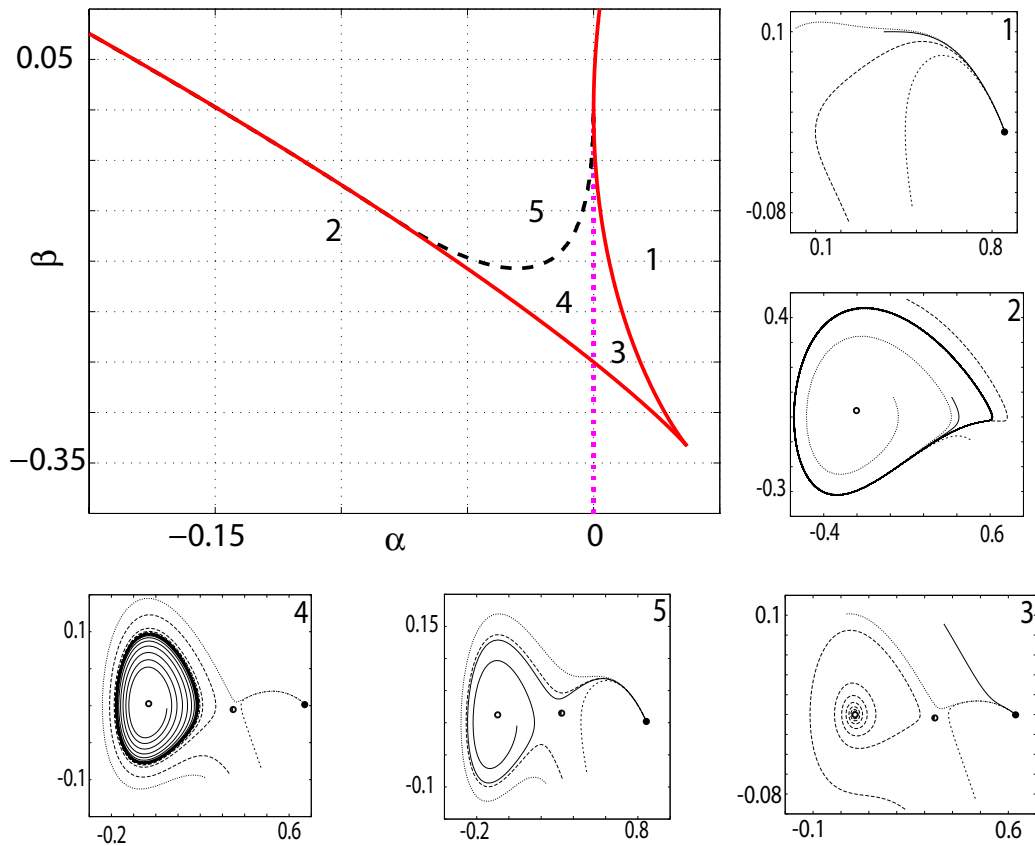


Figura 4.1: Diagrama de bifurcaciones en espacio de parámetros  $(\alpha, \beta)$ , para la forma normal del modelo. En la región con forma de cúspide limitada por la línea sólida, existen tres puntos fijos. En las coordenadas  $\alpha = 0$  y  $\beta = 0$  se observa la colisión tangencial de tres bifurcaciones: Hopf (línea punteada), nodo-silla (línea sólida) y homoclínica (línea rayada) conformando en este punto una bifurcación de *Takens-Bogdanov*. Cuando se cruza de la zona 2 a la zona 5, ocurre una bifurcación SNILC. La integración del modelo en el espacio de fases  $(x, \dot{x})$  con parámetros correspondientes a cada una de la regiones, se muestra en los recuadros correspondientes.

cúbica 4.4, en la que una de ellas es una raíz doble. Para encontrar su ubicación en el espacio de parámetros  $(\alpha, \beta)$ , buscaremos que un extremo de la función sea también una raíz de la misma. Para ello calcularemos e igualaremos a cero la derivada de la ecuación 4.4:

$$\beta + 2x - 3x^2 = 0 \quad (4.5)$$

Luego encontramos sus raíces,

$$x_{1,2} = \frac{1}{3} \pm \sqrt{\frac{1}{9} + \frac{\beta}{3}} \quad (4.6)$$

Se puede observar que para  $\beta > -1/3$  existen dos raíces  $(x_{1,2})$ , cada una de ellas corresponde a uno de los brazos de la bifurcación Cusp. Para parametrizar las curvas de las bifurcaciones, reemplazamos  $x_{1,2}$  en la ecuación 4.4,

$$\alpha = \beta x_{1,2} + x_{1,2}^2 - x_{1,2}^3 \quad (4.7)$$

Las curvas correspondientes a la bifurcaciones nodo-silla, se indican con una curva continua roja en la Figura 4.1.

De modo de identificar el cambio de estabilidad de los puntos fijos, es importante encontrar la ubicación de la bifurcación de Hopf. Para ello, calculamos los autovalores del Jacobiano de las ecuaciones 4.1 y 4.2. Luego, se encuentra el conjunto de parámetros, donde los puntos fijos tengan autovalores complejos con parte real igual a cero, condición necesaria para bifurcación de Hopf (Strogatz, 2001). La bifurcación de Hopf se muestra en la Figura 4.1 con una recta punteada violeta.

Por último, la bifurcación homoclínica, que corresponde a la aniquilación del ciclo límite con el punto silla, fue identificada numéricamente buscando el encendido de oscilaciones dentro de la zona correspondiente a tres puntos fijos. La misma se indica en la Figura 4.1 con una curva rayada negra.

En el diagrama de bifurcaciones completo (Figura 4.1) se enumeran las regiones correspondientes a distintos regímenes dinámicos, del mismo modo que en la Figura 3.8.

Vale recordar que en las regiones 1 y 2, existe un único punto fijo. Dicho punto fijo cambia de estabilidad (cruzando la bifurcación de Hopf) al pasar de la región 1 a la 2 (pasa de estable a inestable). En las zonas 3-5, existen 3 puntos fijos. La zona 3 corresponde a dos puntos fijos estables y un punto silla. Al cruzar la bifurcación de Hopf, se accede a la zona 4. En esta región existe un ciclo límite estable alrededor del punto fijo que cambió su estabilidad. El choque del ciclo límite con el punto silla, define la bifurcación homoclínica y el paso a la región 5. Nuevamente remarcamos que llamamos bifurcación SNILC, a la transición de la región 5 a la región 2.

#### 4.1.2. Escala temporal

Al integrar numéricamente el modelo en una zona cercana a la bifurcación SNILC, observamos que se recupera el comportamiento buscado (una relación inversa entre el contenido espectral y la frecuencia fundamental), pero los tiempos

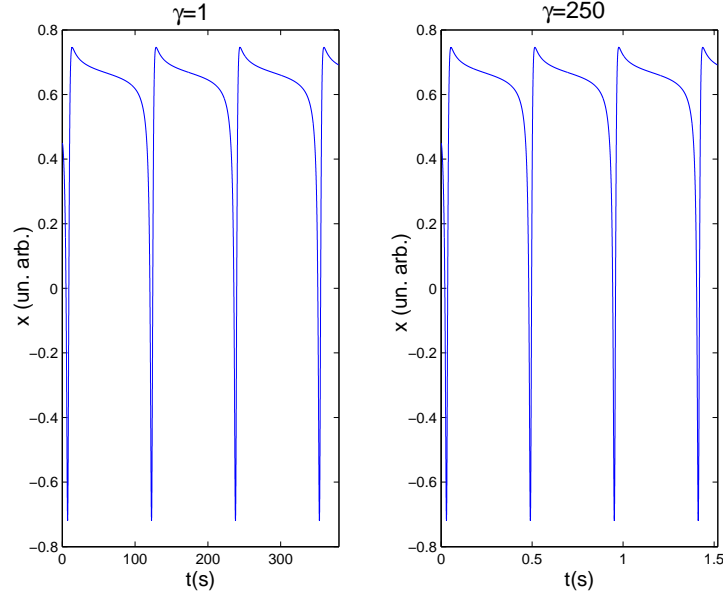


Figura 4.2: Integración numérica para distintas escalas temporales. En papel de la izquierda  $\gamma$  es igual a 1 y en el de la derecha igual a 250. En ambos casos, los parámetros  $\alpha$  y  $\beta$  valen  $-0.149$  y  $0.0$  respectivamente

característicos de las oscilaciones, son demasiado bajos respecto a los valores experimentales observados. Es necesario modificar la escala temporal preservando la dinámica. Para ello, introduciremos un cambio de variable que permita controlar la escala temporal. Definimos una nueva variable temporal  $t = \tau/\gamma$  y reemplazando en las ecuaciones 4.1 y 4.2, obtenemos:

$$\frac{dx}{dt} = \frac{dx}{d\tau} \frac{d\tau}{dt} \quad (4.8)$$

$$= \gamma z \quad (4.9)$$

$$\frac{dz}{dt} = \frac{dz}{d\tau} \frac{d\tau}{dt} \quad (4.10)$$

$$= \gamma(\alpha + \beta x + x^2 - x^3 - xz - x^2 z), \quad (4.11)$$

Luego, haciendo un cambio de variables  $y = \gamma z$  obtenemos:

$$\frac{dx}{dt} = y \quad (4.12)$$

$$\frac{dy}{dt} = \gamma^2 \alpha + \gamma^2 \beta x + \gamma^2 x^2 - \gamma^2 x^3 - \gamma x y - \gamma x^2 y \quad (4.13)$$

Llamaremos a este sistema, “ modelo mínimo ”. En la Figura 4.2 se observan integraciones del modelo en distintas escalas temporales ( $\gamma$  igual 1 y 250) con parámetros  $\alpha$  y  $\beta$  idénticos en ambos casos e iguales a  $-0.149$  y  $0.0$  respectivamente.



El sistema dinámico está diseñando de modo que la dinámica de la variable  $x$ , se comporte de manera similar al desplazamiento desde el equilibrio de la posición de los labios en el modelo físico. Por lo tanto, para generar sonido deberemos integrar el modelo y luego, filtrarlo usando el mismo modelo para el tracto que en el capítulo anterior. En el mismo consideramos la presión al inicio del tracto como,

$$P_i(t) = x(t) - rP_i(t - 2L/v), \quad (4.14)$$

Donde  $L$  corresponde a la longitud del tubo que emula al tracto vocal del ave,  $v$  la velocidad del sonido y  $r$  el coeficiente de reflexión de la onda sonora que arriba al final del tracto vocal.

Con el fin de ajustar el parámetro  $\gamma$  del modelo, se eligen dos segmentos del canto cuyo sonograma se encuentra en el panel superior de la Figura 4.12. Ambos segmentos se encuentran indicados por flechas en la figura. Uno corresponde a una frecuencia fundamental alta, con un comportamiento espectral prácticamente tonal. Mientras que el otro corresponde a una frecuencia fundamental baja, con rico contenido espectral. En ambos casos, se midió la frecuencia fundamental y se calculó el índice de contenido espectral (SCI), del mismo modo que se describió en el capítulo anterior.

Para ajustar el parámetro  $\gamma$ , se procedió del siguiente modo: se tomaron los valores correspondientes al segmento de baja frecuencia fundamental (fundamental = 715 kHz, SCI = 6,6) y se integró el modelo para distintos valores de  $\gamma$  y  $\beta$ . Para cada  $\gamma$ , se buscó el valor de  $\beta$  que correspondiera a la misma frecuencia de oscilación que en el segmento experimental. En dichas coordenadas de  $\gamma$  y  $\beta$ , se calculó el SCI de una integración numérica del sistema. Definimos la distancia  $d_1$  entre el SCI sintético y el experimental como:

$$d_1 = \left( \frac{SCI_{1,sintetico} - SCI_{1,experimental}}{SCI_{1,experimental}} \right)^2 \quad (4.15)$$

Se calculó la distancia en función de  $\gamma$ . Dicha relación se encuentra graficada en la Figura 4.3 con una curva azul continua. Se repitió el mismo procedimiento con el segmento correspondiente a sonidos de alta frecuencia fundamental (fundamental = 5468 Hz, SCI = 1,0245), llamando a la distancia obtenida  $d_2$  (se graficó en la misma figura con una curva punteada negra). Combinamos ambas distancias de forma  $d_{total} = d_1 + d_2$ , la cual tiene un mínimo para  $\gamma \sim 23500$ . La distancia combinada se graficó también en la misma figura con una curva roja rayada.

Intuitivamente, el contenido espectral de una señal sintética integrada con parámetros  $(\alpha, \beta)$ , depende básicamente de la distancia entre estos parámetros y la curva en el espacio de parámetros que indica la bifurcación SNILC. Por otro lado, la frecuencia fundamental de esa misma integración dependerá fuertemente del valor del  $\gamma$ , ya que el mismo define la escala temporal. De ese modo, hemos ajustado  $\gamma$  para obtener una señal con la misma relación entre la frecuencia fundamental y el

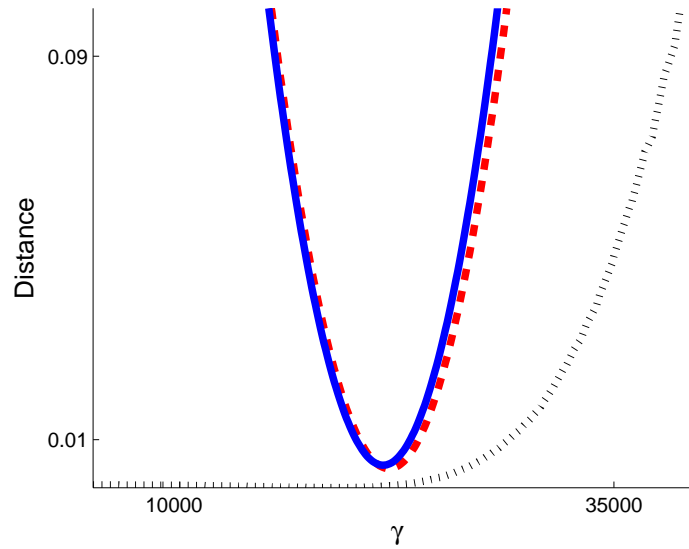


Figura 4.3: Ajuste del parámetro  $\gamma$ . Distancia entre los valores experimentales de contenido espectral y sus pares teóricos en función del parámetro  $\gamma$ . La línea negra punteada corresponde al segmento de alta frecuencia fundamental; la curva rayada roja al segmento de baja frecuencia fundamental y la línea azul continua a la combinación de estas dos. La curva de la distancia combinada, alcanza un mínimo para un valor aproximado de  $\gamma \sim 23500$ .

SCI que la observada en las grabaciones experimentales del canto. En la Figura 4.4 se muestra el diagrama de bifurcaciones del sistema en el espacio  $(\alpha, \beta)$ . Se indican en la misma figura, curvas de iso-frecuencia (integraciones del sistema para distintos  $(\alpha, \beta)$  con valores idénticos de frecuencia fundamental) e iso-SCI (ídem anterior pero con idénticos SCI) para parámetros  $\gamma = 23500$ ,  $L = 1,9 \text{ cm}$  y  $r = 0,9$ .

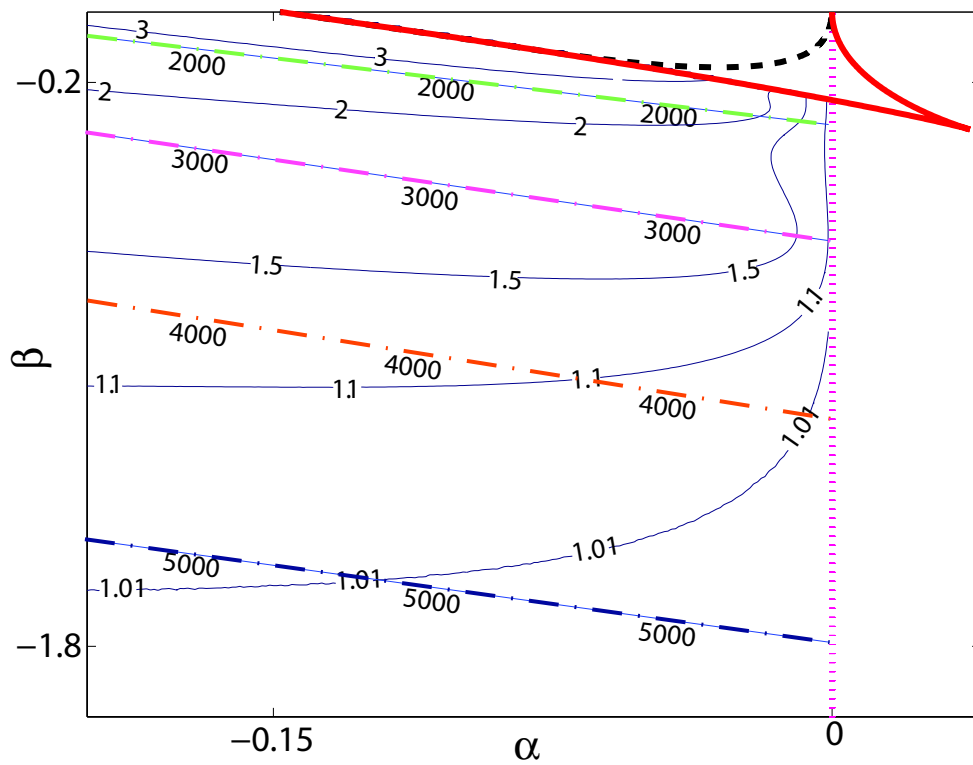


Figura 4.4: Vista detallada del diagrama de bifurcaciones en el rango de parámetros usado en la generación sintética de canto de Diamante Mandarin. Las curvas correspondientes a las bifurcaciones, se encuentran indicadas del siguiente modo: Nodo-silla, continua roja; Hopf, punteada violeta y Homoclínica, rayada negra. Debajo de la bifurcación Nodo-silla, se encuentran curvas correspondientes a integraciones del sistema. Las mismas se encuentran rotuladas con sus valores de SCI (líneas sólidas) y frecuencia fundamental (líneas rayadas). El diagrama fue realizado con los parámetros  $\gamma = 23500$ ,  $L = 1,9 \text{ cm}$  y  $r = 0,9$ .

## 4.2. Siringe analógica

En una primera instancia, se desarrolló una solución analógica para integrar en tiempo real el modelo planteado. En este primer abordaje, la intención fue sintetizar las sílabas de baja frecuencia. El ave regularía la emisión de dichas sílabas controlando únicamente la presión en los sacos aéreos. Para ello desarrollaremos un dispositivo que utilice únicamente como parámetro de entrada al sistema, la presión en los sacos aéreos. Por lo que, en esta instancia, se considero únicamente el parámetro  $\alpha$  tomando  $\beta = 0$ . De ese modo, la ecuación a integrar es:

$$\frac{dx}{dt} = y \quad (4.16)$$

$$\frac{dy}{dt} = \gamma^2\alpha + \gamma^2x^2 - \gamma^2x^3 - \gamma xy - \gamma x^2y \quad (4.17)$$

Este integrador analógico se basa en la técnica previamente usada en la literatura. Dicha técnica se basa en la utilización de amplificadores operacionales realimentados y multiplicadores (Zysman et al., 2005).

Se utilizaron amplificadores operacionales TL081 (Texas Instruments, Single OPAMP) y TL082 (Texas Instruments, Dual OPAMP). Para multiplicar señales eléctricas se utilizaron multiplicadores analógicos AD633JN (Analog Devices). El esquema de dichos multiplicadores se encuentra en la Figura 4.5 y su salida está dada por:

$$W = \frac{(X1 - X2)(Y1 - Y2)}{10V} - Z, \quad (4.18)$$

En este caso utilizaremos  $X2 = Y2 = Z = 0$  y en el caso de necesitar calcular el cuadrado de una señal,  $X1 = Y1$ .

El esquema completo del integrador analógico de las ecuaciones diferenciales del modelo, se presenta en la Figura 4.6.

El integrador 1 (I1), conformado por un amplificador operacional con realimentación negativa mediante un capacitor (Horowitz and Hill, 1989), relaciona los puntos  $v$  y  $w$  del siguiente modo:

$$v = -\frac{1}{RC} \int w dt \quad (4.19)$$

Los cuatro multiplicadores calculan los términos  $v^2, vw, v^2w$  y  $v^3$ . La señal de este último luego pasa por un inversor, desarrollado con un amplificador operacional con realimentación negativa mediante un cable (Horowitz and Hill, 1989). El integrador 2 (I2) suma las corrientes aportadas por los multiplicadores y la señal de entrada ( $\alpha$ ). La salida del mismo se calcula como:

$$w = -\frac{1}{C} \int \left( \frac{\alpha}{R_1} + \frac{v^2}{10R_2} + \frac{vw}{10R_3} - \frac{v^3}{100R_4} + \frac{v^2w}{100R_5} \right) dt \quad (4.20)$$

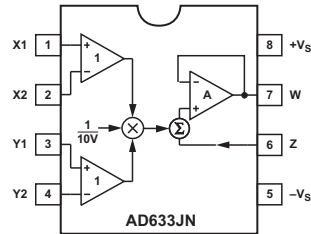


Figura 4.5: Multiplicador AD633

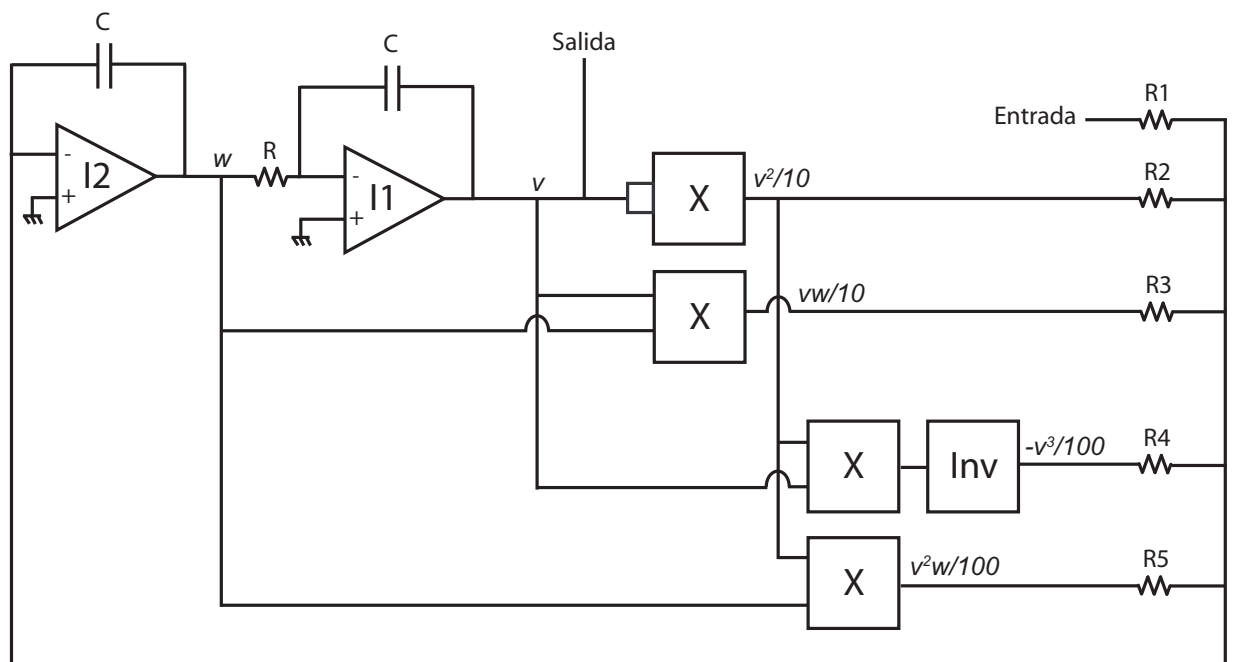


Figura 4.6: Esquema del integrador analógico.

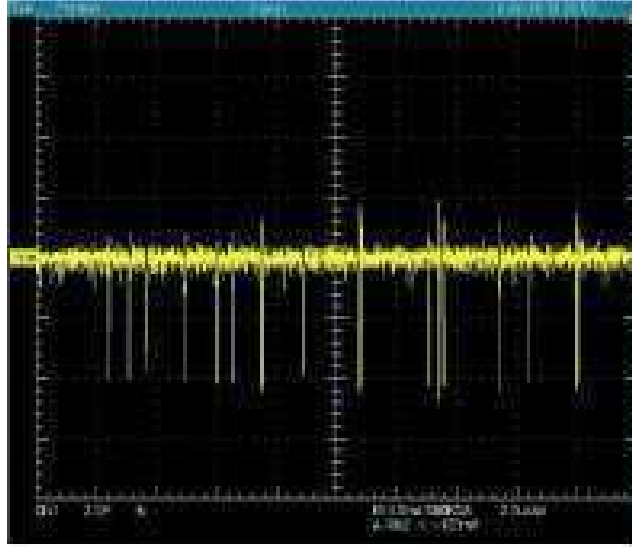


Figura 4.7: Siringe analógica, oscilaciones de frecuencia no definida para un parámetro de control cercano a la bifurcación.

Reemplazando  $x = v$  e  $y = -\frac{w}{RC}$ , obtenemos la ecuación diferencial integrada por el circuito:

$$\frac{dx}{dt} = y \quad (4.21)$$

$$\frac{dy}{dt} = \frac{\alpha}{R_1 RC^2} + \frac{x^2}{10R_2 RC^2} - \frac{xy}{10R_3 C} - \frac{x^3}{100R_4 RC^2} - \frac{x^2 y}{100R_5} \quad (4.22)$$

Si ajustamos el valor de los componentes de la siguiente manera:  $R_1 = \frac{1}{\gamma^2 RC^2}$ ,  $R_2 = \frac{1}{10\gamma^2 RC^2}$ ,  $R_3 = \frac{1}{10\gamma C}$ ,  $R_4 = \frac{1}{100\gamma^2 RC^2}$  y  $R_5 = \frac{1}{100\gamma C}$ , el sistema resulta equivalente al propuesto en las ecuaciones 4.16 y 4.17.

Un problema que se encontró con este diseño, es que las señales correspondientes a  $v$  y  $w$  son de muy bajo nivel de tensión, siendo comparables con el ruido eléctrico esperable en un montaje de las características implementadas. Esta situación afecta la estabilidad en frecuencia de las oscilaciones. En la Figura 4.7 se muestra la salida del sistema para un parámetro de control cercano al valor de bifurcación. Como se puede observar, las oscilaciones no tienen una frecuencia definida.

La solución encontrada para esta situación, fue realizar una amplificación de dichas variables mediante un cambio en la escala de amplitud. Para ello, haremos un nuevo cambio de variables en el sistema  $x_* = \eta x$  e  $y_* = \eta y$ . Donde  $\eta$  representa el factor de escala de amplitud. De ese modo el sistema queda:

$$\frac{dx_*}{dt} = y_* \quad (4.23)$$

$$\frac{dy_*}{dt} = \eta\gamma^2\alpha + \frac{\gamma^2x_*^2}{\eta} - \frac{\gamma^2x_*^3}{\eta^2} - \frac{\gamma x_* y_*}{\eta} - \frac{\gamma x_*^2 y_*}{\eta^2} \quad (4.24)$$

Modificando el valor de  $\eta$  se logra amplificar las señales correspondientes a las variables del modelo. Por ende, se logra que la integración sea menos sensible al ruido eléctrico. En la Figura 4.8 puede observarse una foto del montaje definitivo de la siringe analógica. Los valores utilizados fueron:  $R = 47 K\Omega$ ,  $R_1 = 4.7 K\Omega$ ,  $R_2 = 10 K\Omega$ ,  $R_3 = 4.7 K\Omega$ ,  $R_4 = 4.7 K\Omega$ ,  $R_5 = 2.2 K\Omega$ ,  $C = 1 nF$ .

Para poder sintetizar sonido similar al producido por el ave, fue necesario conectar la salida del integrador analógico a un filtro electrónico analógico que represente el filtrado producido por el tracto vocal del ave. Para ello, se calcularon las características de transferencia del tracto modelado (ver 2.6 y 2.7). Luego se implementó un filtro pasabanda biquad (Horowitz and Hill, 1989), con el pico de transferencia para  $f_0 = 4200 Hz$  y ancho de banda  $BW = 1000 Hz$ .

En la Figura 4.9 se puede observar la relación entre el contenido espectral (SCI) y la frecuencia fundamental, tanto para la siringe electrónica analógica implementada como para el modelo propuesto. Como puede observarse, el sistema reproduce las características acústicas buscadas.

Mediante el desarrollo de la siringe electrónica analógica se lograron sintetizar algunas de las sílabas presentes en el canto del Diamante Mandarín. En particular, se pudieron sintetizar las sílabas de baja frecuencia fundamental y alto contenido espectral, controlando el parámetro asociado a la presión en los sacos aéreos ( $\alpha$ ) sin modificar el parámetro relacionado a la actividad de los músculos ventrales ( $\beta$ ). Dicha estrategia es consistente con datos experimentales que no encuentran actividad en los músculos siringeos ventrales en dichas sílabas (Vicario, 1991).

Se encontraron limitaciones en el enfoque tecnológico de integración analógica. Las principales están dadas, porque los parámetros del sistema dependen de componentes físicos (x ej, resistencias, capacitores, transistores) y por ello es difícil ajustar con precisión sus valores o cambiar los mismos ante modificaciones en el modelo. También es importante indicar la dependencia de los mismos con factores externos como la temperatura.

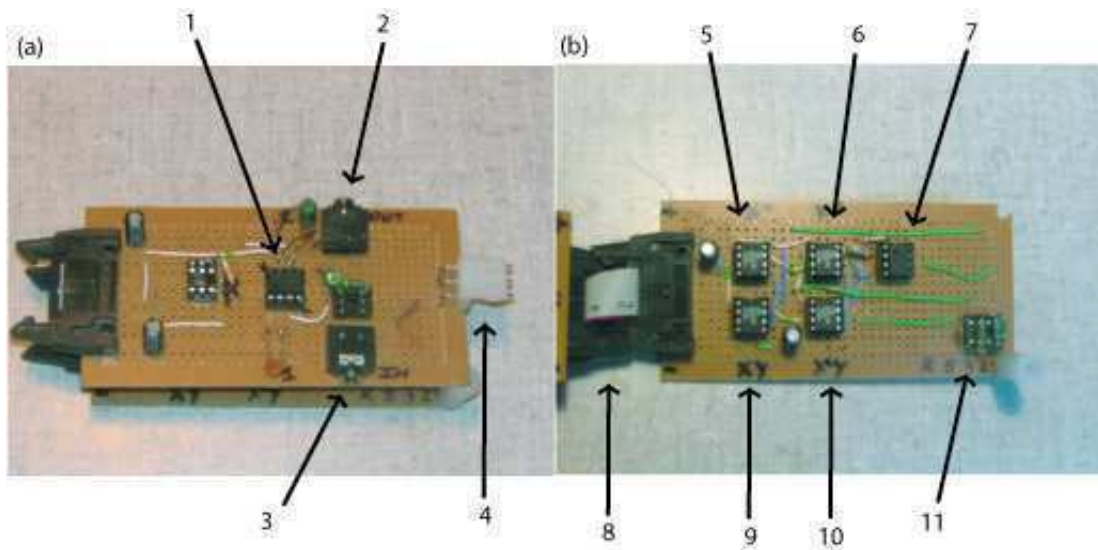


Figura 4.8: Fotos de la siringe analógica. (a) Capa superior de la siringe analógica: (1) Doble amplificador operacional configurado con dos integradores. (2) Conector señal de salida. (3) Conector señal de entrada. (4) Conector de alimentación. (b) Capa inferior de la siringe analógica: (5) Multiplicador término  $x^2$ . (6) Multiplicador  $x^3$ . (7) Inversor para el término  $x^3$ . (8) Cable plano conector entre capas. (9) Multiplicador término  $xy$ . (10) Multiplicador término  $x^2y$ . (11) Resistencias de control de los parámetros.

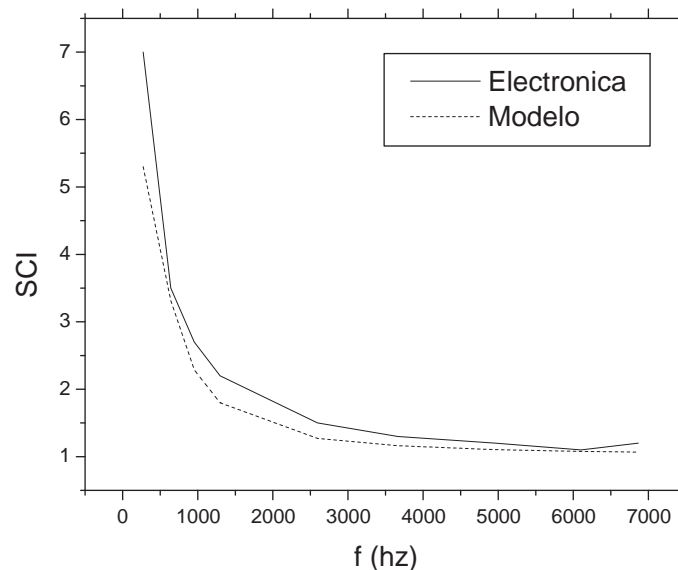


Figura 4.9: Relación SCI vs. frecuencia fundamental. En la curva continua se observa la relación para la siringe electrónica analógica. En la curva rayada, para el modelo planteado.



### 4.3. Siringe digital

El siguiente paso a seguir fue desarrollar una nueva versión de la siringe electrónica, que incluya la capacidad de controlar el parámetro correspondiente a la actividad de los músculos siringeos ventrales ( $\beta$ ). Para desarrollar dicho dispositivo, se planteó un cambio en la tecnología, reemplazando la integración analógica del modelo por un sistema que integre el modelo de forma digital. Buscando de ese modo salvar las limitaciones encontradas durante el enfoque previo.

Una solución digital del problema, permite ajustar con sencillez y precisión los parámetros del modelo a integrar. Por otro lado, dado que se busca desarrollar el proceso en tiempo real, la solución digital presenta un gran desafío. Como la salida del sistema es sonido con contenido espectral de hasta aproximadamente 12  $KHz$  es necesario, por el teorema de Nyquist (Smith, 1999), que la frecuencia de muestreo de salida sea como mínimo 24  $KHz$ . En el lapso de tiempo entre dos muestreos de salida, el equipamiento debe integrar numéricamente el modelo. Simultáneamente, el sistema debe adquirir las señales de entrada que utilizará para modificar los parámetros del modelo. El muestreo de las señales de entrada debe ser de una frecuencia menor o igual a la frecuencia de muestreo de salida.

Es importante indicar que el paso de integración debe ser lo suficientemente chico como para preservar la precisión de dicha integración. Para que la integración no se adelante ni se atrase respecto a los sucesos externos, es necesario que el tiempo integrado entre muestreos sea exactamente igual al transcurrido entre los mismos (por ejemplo, si la frecuencia de muestreo es de 20  $KHz$ , el tiempo entre muestreos es de 50  $\mu s$ , entonces el sistema podrá integrar dicho tiempo en 20 iteraciones de paso con integración de 2.5  $\mu s$ ).

#### 4.3.1. ¿Qué es un DSP?

Un procesador digital de señales o DSP (*digital signal processor*) es un dispositivo programable, diseñado para acelerar y facilitar la programación de algoritmos de: muestreo, filtrado, conversión, codificación, detección, etc. de señales digitales (Chassaing, 2002). Dichos dispositivos están presentes en la mayoría de equipos electrónicos actuales (por ej, celulares, PDAs). El DSP es generalmente acompañado por conversores, de modo que el proceso de la señal siga una secuencia como la que se muestra en el esquema de la Figura 4.10. Internamente los DSP se encuentran especializados para realizar cálculos matemáticos. Para ello utilizan estructuras especializadas como: multiplicadores-sumadores (MACs), buffers circulares y arquitectura Harvard para separar la memoria del programa de la de los datos (Smith, 1999).

Los DSP pueden ser clasificados en dos categorías, de punto fijo y de punto flotante. Esta clasificación se refiere al formato usado para manipular y guardar números dentro del dispositivo. Los DSP de punto fijo típicamente representan los números como enteros sin signo de 16 bits. En comparación, la representación en



Figura 4.10: Esquema del procesamiento digital de señales. El mismo sigue una secuencia de conversión analógica-digital de una señal de entrada, manipulación digital en el DSP y conversión digital-analógica de la señal de salida.

punto flotante utiliza 32 bits y a diferencia de la de punto fijo, la representación de los números no está uniformemente espaciada y se realiza utilizando notación científica (Press et al., 1992). Los DSP de punto flotante tienen mejor precisión, mayor rango dinámico y su programación es más sencilla. Dado que se busca desarrollar un integrador numérico del modelo matemático, para lo cual es importante la representación precisa de números con diferentes órdenes de magnitud, se optó por utilizar un DSP punto flotante.

Los DSP pueden ser programados usualmente en assembler o en C. Los programas escritos en assembler son ejecutados más rápidamente. Por otro lado, los escritos en C son más fáciles de desarrollar y mantener. En nuestro caso, el programa se desarrolló en C de modo de poder ajustar fácilmente los parámetros del modelo. Se deberá evaluar a futuro el límite de velocidad de ejecución, implementando parte o la totalidad del programa en assembler.

#### 4.4. Desarrollo de canto sintético

Una de las características más interesantes del canto del Diamante Mandarín es la variabilidad de la frecuencia fundamental dentro de cada sílaba. Por lo tanto, para sintetizar canto, deberemos integrar las ecuaciones que determinan la dinámica de los labios con parámetros que varíen en el tiempo. Dichos parámetros corresponderán a las instrucciones fisiológicas que reciben los músculos que controlan la siringe para generar el canto.

Para ello, utilizaremos como guía la Figura 4.4 para poder seleccionar los caminos en el espacio de parámetros que permiten sintetizar un canto de características acústicas temporales similares. Para asegurar la estabilidad de la integración, es importante que el movimiento en el espacio de parámetros sea mucho más lento que la dinámica de los labios. Esta aproximación puede considerarse válida dado que las sílabas más cortas producidas por estas aves son de aproximadamente 100 ms, mientras que las vocalizaciones de más baja frecuencia son del orden de cientos de Hz.

Como describimos en el Capítulo 2, las mediciones en el flujo de aire permitieron determinar la contribución relativa de cada una de las fuentes de la siringe en la generación de canto en el Diamante Mandarín (Goller and Cooper, 2004). Recordamos que básicamente las sílabas de alta frecuencia son generadas en la fuente derecha de la siringe. Las de baja frecuencia, típicamente involucran la activación de ambas fuentes durante parte o la totalidad de la sílaba. Para estas sílabas de baja frecuencia, cuando ambas fuentes están activas se han reportado oscilaciones pulsátiles en fase y en contra fase (Jensen et al., 2007).

De modo de encontrar los trayectos en el espacio de parámetros necesarios para producir canto sintético, definiremos dos regímenes. Asumiremos que los sonidos de alta frecuencia (frecuencia fundamental mayor a 1,8kHz) se encienden mediante una bifurcación de Hopf, generando de ese modo oscilaciones tonales. Mientras que para los sonidos de baja frecuencia (frecuencia fundamental menor a 1,2kHz), asumiremos que comienzan al cruzar la bifurcación SNILC. Para estas vocalizaciones consideraremos, en el caso de que no haya discontinuidades espectrales, que ambas fuentes de la siringe están activas y operando en el mismo régimen dinámico.

Se construyeron dos series temporales del parámetro  $\beta$  del modelo. Cada  $\beta_{1,2}^{numerico}(t)$  corresponde a uno de los regímenes mencionados. Cada serie fue definida de modo que sinteticen sonidos con secuencias de frecuencia fundamental que ajusten el canto observado. Cada  $\beta_{1,2}^{numerico}(t)$  fue usado en la integración del modelo, generando dos series temporales independientes  $x_{1,2}$ . Se generó el canto sintético sumando ambas  $x$  emulando la presión al inicio del tracto,

$$P_i(t) = x_1(t) + x_2(t) - r P_i(t - 2L/v). \quad (4.25)$$

En la Figura 4.11 (c,d) se muestran los valores de  $\beta_{1,2}^{numerico}(t)$  utilizados para sintetizar el canto que se muestra en el panel del medio de la Figura 4.12. En cuanto al parámetro  $\alpha$ , se eligieron dos valores  $\alpha_{on} = -0,15$  y  $\alpha_{off} = 0,05$  que encendieran

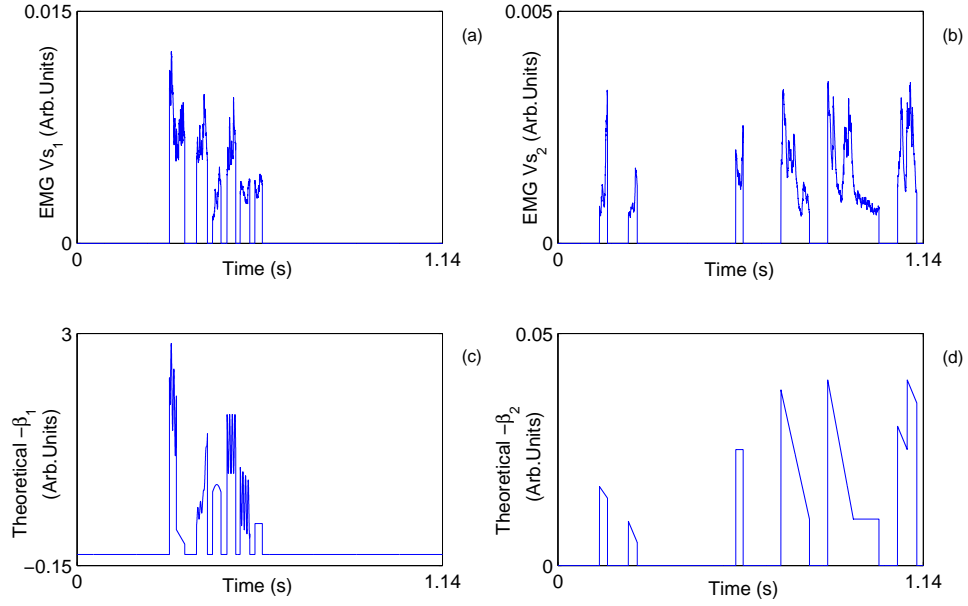


Figura 4.11: Actividad electromiografica típica del músculo vS y serie temporal de  $\beta$ . (a) Actividad del músculo vS derecho durante los intervalos de tiempo en el que se emitieron sílabas de alta frecuencia. (b) Actividad del músculo vS izquierdo durante los intervalos de tiempo en el que se emitieron sílabas de baja frecuencia. (c,d)  $-\beta_{1,2}^{numerico}(t)$  definido de modo que la frecuencia fundamental sintetizada corresponda con la del canto observado.

o apagarán las oscilaciones. La transición entre estos dos estados, es determinada ante la detección de una presión en los sacos aéreos mayor a un valor umbral y la aparición de vocalizaciones sonoras.

De modo de cuantificar la similitud entre el canto natural y el sintético, se midieron frecuencia fundamental y contenido espectral en cada caso. La frecuencia fundamental fue extraída cada 10 *ms* usando el método descrito en (Boersma and Weenink, 2001). De ese modo, se obtuvieron series temporales de la frecuencia fundamental para el canto natural y el sintético. Definimos distancia de frecuencia fundamental del siguiente modo:

$$D_{ff} = \left( \frac{ff_1 - ff_2}{ff_1/2 + ff_2/2} \right)^2, \quad (4.26)$$

Donde  $ff_1$  es la serie temporal de frecuencia fundamental del canto experimental y  $ff_2$  es la serie temporal de frecuencia fundamental del canto sintético. La distancia media entre el canto experimental y el sintético fue de 0,082 con desviación estándar (*STD*) de 0,209.

Se calculó el SCI en una ventana temporal de 2.7 *ms*, alrededor de cada uno de los lugares donde la frecuencia fundamental fue medida. Definimos la distancia *SCI*

( $D_{SCI}$ ) como:

$$D_{SCI} = \left( \frac{SCI_1 - SCI_2}{SCI_1/2 + SCI_2/2} \right)^2, \quad (4.27)$$

Donde  $SCI_1$  corresponde al canto natural y  $SCI_2$  al canto sintético. La distancia media entre el canto natural y el sintético es de 0,134 con  $STD$  igual a 0,285.

Una vez caracterizadas las series temporales de los parámetros necesarios para la síntesis de canto sintético con frecuencia fundamental y contenido espectral similar al canto natural, se procedió a realizar dos pasos. Primero, se encontró una transformación capaz de mapear la actividad electromiográfica de los músculos siringeos de un pájaro cantando, a los parámetros computados anteriormente. Y segundo, se implementó dicha transformación y el modelo en el DSP.

Encontrar la transformación adecuada de las señales de actividad electromiográfica (EMG) de los músculos siringeos a los parámetros dependientes del tiempo, permite generar el puente necesario para poder sintetizar canto comandado por instrucciones fisiológicas. La actividad de los músculos siringeos se registró insertando electrodos bipolares (California Fine Wire, acero inoxidable, 25  $\mu m$ , cubiertos por aislante excepto en el extremo) en los músculos. Los mismos fueron fijados a los músculos usando una microgota de pegamento y direccionados hacia una mochila dispuesta en la espalda, donde se conectan a un pre-amplificador. Los electrodos fueron insertados en ambos músculos *Siringialis ventralis* ( $vS$ ). La actividad de dichos músculos, como se indicó en el Capítulo 2, por lo menos en otras especies de oscinos, está fuertemente correlacionada con la frecuencia de las oscilaciones de los labios (Larsen and Goller, 1999).

La señal electromiográfica se amplifica (amplificador DAGAN EX4-400) y se filtra con un pasa banda entre 100 y 3000  $Hz$ . Luego se procede a rectificar y suavizar la señal de EMG adquirida. Dicho procedimiento, provee de una cuantificación de la actividad muscular. Esta metodología fue usada previamente para entender cómo los músculos contribuyen a los movimientos responsables de la evolución lenta de la frecuencia fundamental dentro de las sílabas del canto (Mindlin et al., 2003). Este preprocesamiento del EMG fue realizado mediante un circuito que integra la siguiente ecuación:

$$\frac{dvS}{dt} = -\tau vS + |vS_{EMG}| \quad (4.28)$$

Donde  $|vS_{EMG}|$  es el valor absoluto de la actividad electromiográfica medida y  $vS$  es la actividad suavizada.

Grabaciones de la actividad de  $vS$  son ilustradas en la Figura 4.11(a,b). En la Figura 4.11(a) se muestra un segmento de actividad en el  $vS$  derecho, en el intervalo de tiempo en que sonidos de alta frecuencia son emitidos y en la Figura 4.11(b) se muestra el  $vS$  izquierdo, en el intervalo de tiempo en que sonidos de baja frecuencia son emitidos (en cual tanto la fuente izquierda y la derecha estarían activas).

De modo de seleccionar el conjunto óptimo de señales fisiológicas para conducir el modelo, se graficó la señal del  $vS$  en función de la frecuencia fundamental. Dado que

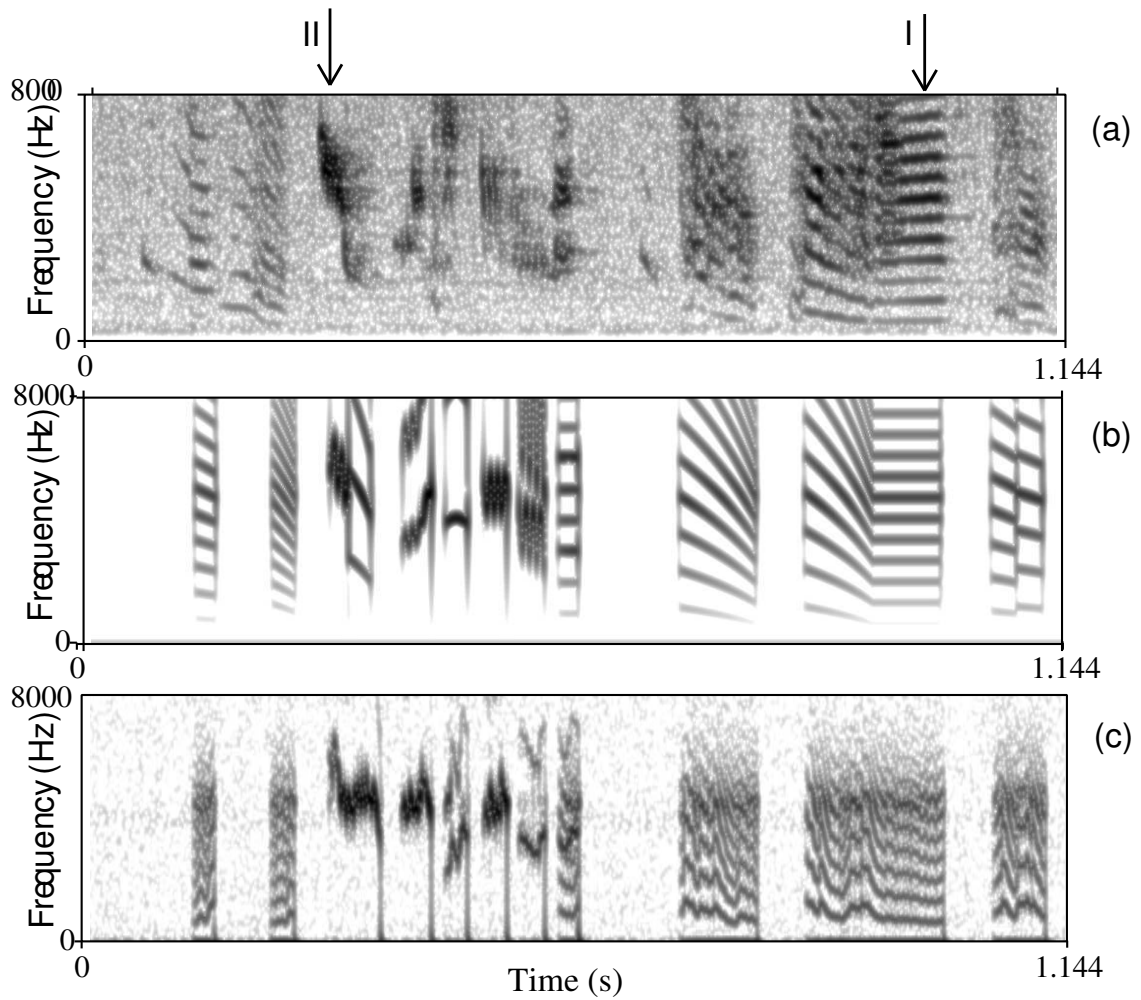


Figura 4.12: (a) Sonograma de un canto típico de Diamante Mandarin. Las flechas indican los segmentos utilizados para ajustar el parámetro  $\gamma$ . (b) Canto sintético utilizando  $\beta_{1,2}^{numerico}(t)$  para conducir el modelo. (c) Canto sintetizado en tiempo real, conduciendo el modelo implementado en el DSP con los registros de actividad fisiológica.

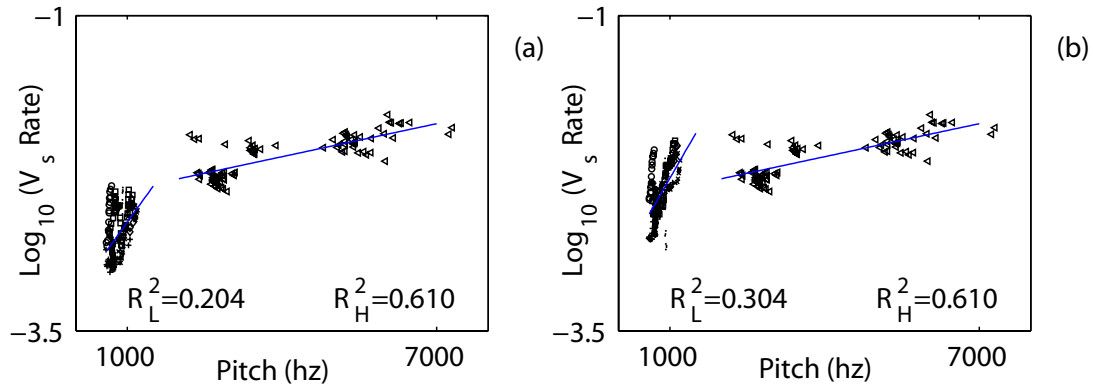


Figura 4.13: Actividad del músculo  $vS$  en función de la frecuencia fundamental. En el panel de la izquierda (a), la frecuencia fundamental para las sílabas de baja frecuencia (fundamental  $< 1,8 \text{ kHz}$ ) fue graficada en función de la actividad del músculo ventral izquierdo. En (b), la fundamental de las mismas sílabas fue graficada como función de la actividad del músculo derecho. En ambos paneles, las sílabas de alta frecuencia fueron graficadas usando la actividad del músculo derecho.

se ha establecido que únicamente la parte derecha de la siringe es responsable de las vocalizaciones de alta frecuencia (Goller and Cooper, 2004), es natural seleccionar la actividad del músculo  $vS$  derecho para generar dicha parte de la relación. Por lo tanto, todas las vocalizaciones de frecuencia fundamental mayor a  $1,8 \text{ kHz}$  fueron graficadas usando la actividad del músculo  $vS$  derecho de la siringe en función de la frecuencia fundamental.

Dado que ambas fuentes están activas en las sílabas de baja frecuencia, el modo de tratarlas es más sutil. Hemos seleccionado la actividad del músculo  $vS$  izquierdo como parámetro para conducir el modelo. Se hizo esta selección luego de inspeccionar las actividades izquierda y derecha, en función de la frecuencia fundamental. En el panel izquierdo de la Figura 4.13 se muestra, para sonidos de alta frecuencia, la actividad derecha en función de la frecuencia fundamental y para los sonidos de baja frecuencia, la actividad izquierda en función de la frecuencia fundamental. En el panel derecho de la misma figura, se utiliza la actividad derecha en todas las sílabas. En dicho panel, se puede verificar que hay niveles de actividad con dos posibles frecuencias fundamentales asociadas. Esto indicaría que tanto un músculo no medido como un acople con el lado izquierdo, son necesarios para diferenciar en que régimen opera la siringe. La única combinación que define la relación biyectiva necesaria, es cuando la actividad izquierda es usada para vocalizaciones de baja frecuencia y la actividad derecha para los sonidos de alta frecuencia; por ende, es esta la combinación óptima para conducir el modelo.

De modo de sintetizar canto al conducir el modelo representado por (4.12,4.13) con instrucciones fisiológicas, es necesario encontrar una transformación desde las mediciones electromiograficas hacia los parámetros del modelo  $(\alpha, \beta)$ . Como antes,  $\alpha$  tomará únicamente uno de los dos valores posibles representando el gating de sistema. Mientras que  $\beta$  será responsable de las modulaciones de frecuencia y contenido espectral de los sonidos sintetizados.

Para encontrar la transformación deseada, se propone un  $\beta^{propuesto}$  tal que:

$$\beta_i^{propuesto} = a_{0,i} + a_{1,i}vS_i(t) + a_{2,i}vS_i(t)^2 \quad (4.29)$$

Donde los parámetros  $a_i$  son elegidos de modo que  $\chi^2 = \chi^2(\beta_i^{propuesto} - \beta_i^{numerico})$  sea mínimo, donde  $i = 1, 2$  corresponden a los regímenes de alta y baja frecuencia y a los registros electromiograficos derecho e izquierdo respectivamente. En nuestro ejemplo, los valores obtenidos son  $(a_{0,1}, a_{1,1}, a_{2,1}) = (-6697, 152,65, 0,0848)$  y  $(a_{0,2}, a_{1,2}, a_{2,2}) = (0, -17,79, 0,0016)$ .

Habiendo hallado esta transformación, tenemos las herramientas para conducir fisiológicamente un dispositivo que integre las ecuaciones del modelo.



## 4.5. Instrucciones fisiológicas conduciendo el DSP

En enfoques previos para la construcción de dispositivos biomiméticos que integren en tiempo real, las estrategias propuestas se basaron en la construcción de circuitos electrónicos analógicos (Sitt and Aliaga, 2007; Zysman et al., 2005). Estos dispositivos, como se indicó previamente, tienen el inconveniente de que los parámetros son difíciles de ajustar y dependen de aspectos externos (por ejemplo, la temperatura ambiente). Es por ello que en este caso utilizamos un procesador digital de señales (DSP). El mismo forma parte de en una placa de desarrollo inicial (DSK), que lo provee de las conexiones necesarias de entrada/salida (Chassaing, 2002). Esta placa de desarrollo permite adquirir una señal, usar los datos adquiridos como entrada en un programa previamente cargado y enviar los resultados hacia el exterior. Por ende, la implementación del modelo requiere solo la programación de un código que integre el modelo eficientemente.

En este caso se eligió la placa de desarrollo *Texas Instruments TMS320C6713 DSK*, que básicamente consiste de un DSP TI6713 operando a 225 MHz y conversores analógico-digital y digital-analógico (Chassaing, 2002).

Se escribió un programa en C de modo que el mismo realice los siguientes pasos:

1. Utilizar la capacidad de adquisición de la placa para digitalizar las señales correspondientes a los registros fisiológicos.
2. Mediante la transformación previamente encontrada, calcular los parámetros  $(\alpha, \beta)$  a partir de la señales adquiridas.
3. Integrar numéricamente el modelo.
4. Filtrar la salida de la integración usando el modelo del tracto.
5. Convertir la salida del filtro en una señal analógica (canto sintético).

La placa digitalizó las señales de entrada con una frecuencia de muestreo de 48 KHz. En el lapso de tiempo entre dos muestreos, el procesador integra las ecuaciones correspondientes a la actividad de los labios y procesa el filtrado pasivo del tracto. Como método numérico de integración, se eligió un Forward Euler de paso fijo (Press et al., 1992) dado su bajo costo computacional (Paja et al., 2007). El DSP realiza 18 iteraciones del integrador en el tiempo entre muestreos (cada iteración tiene un tiempo de integración  $\Delta t = \frac{1/(48kHz)}{18}$ , de modo de mantener la sincronía).

Se realizan dos integraciones independientes del modelo representado por el sistema (4.13). Cada una corresponderá a uno de los regímenes descriptos. Las soluciones fueron sumadas como presiones al inicio del tracto de siguiente modo:

$$P_{entrada}(t) = x_1(t) + x_2(t) - rP_{entrada}(t - 2L/v) \quad (4.30)$$

La salida del filtro fue convertida a una señal analógica por el DSP a la misma frecuencia en que la entrada fue digitalizada. El sonograma de la señal de salida se muestra en el panel inferior de la Figura 4.12.

Más allá de la similitud cualitativa, se cuantificó la similitud entre el canto natural y el sintetizado por el DSP. Se usó la metodología anteriormente descrita. En caso de la frecuencia fundamental, la distancia es  $D_{ff} = 0,189$  con desviación estándar  $STD = 0,169$ , mientras que para el SCI,  $D_{SCI} = 0,135$  y  $STD = 0,215$ .

En este capítulo analizamos la construcción de un dispositivo biomimético electrónico capaz de transformar instrucciones fisiológicas en canto. La solución óptima encontrada, fue programar un procesador digital de señales (DSP) para que lea los datos experimentales de entrada, integre las ecuaciones dinámicas correspondientes a una versión minimal del modelo físico de la siringe y use las soluciones para sintetizar canto.

La generación en tiempo real de sonidos de alta frecuencia con contenido espectral específico representa un desafío. En este capítulo mostramos que el estado actual de la tecnología puede superar este desafío. El uso de tecnología digital representa una mejora cualitativa respecto de los enfoques previos de soluciones biomiméticas vocales (Zysman et al., 2005): el control de los parámetros del sistema, ya sea para mejoras o como herramienta para desarrollar experimentos de feedback auditivo alterado, resulta flexible y precisa.



# Conclusiones

---

En las últimas décadas, el canto de las aves oscinas ha servido como modelo animal para entender uno de los efectos más fascinantes que puede presentar un sistema nervioso: la capacidad de modificarse con el fin de cumplir una tarea específica mediante el ajuste de las conectividades entre las neuronas que lo constituyen. La posibilidad de tener un sistema animal capaz de aprender un comportamiento tan complejo como la vocalización, ha inspirado decenas de esfuerzos que cubren desde la etología, la neurociencia, la fisiología hasta, como en esta tesis, la física.

Entre las experiencias que más pueden aportar a la dilucidación de este proceso, se encuentran las experiencias de retroalimentación acústica alterada. En éstas, un ave es expuesta a un canto distinto del esperado y realiza ajustes compensatorios cuya naturaleza permitiría comprender los códigos de corrección de errores, o de refuerzo de aciertos, que se encuentran en el corazón del proceso de aprendizaje. Esta tesis se inscribe en ese programa de investigación, y ha guiado distintos esfuerzos a lo largo de los últimos años en el Laboratorio de Sistemas dinámicos: ¿es posible construir un dispositivo biomimético capaz de generar canto sintético controlado por un ave?

Este programa es incremental por naturaleza. Para lograr este objetivo, hay que cumplir una enorme cantidad de estudios intermedios. La estrategia, desde un comienzo, fue elaborar modelos matemáticos que describieran los mecanismos físicos involucrados en la generación del canto, implementarlos electrónicamente, así como desarrollar las técnicas de medición de las variables fisiológicas que han de, finalmente, controlar el dispositivo. Cada una de estas etapas involucró el modelado, la elección de las tecnologías pertinentes, así como la implementación. En el camino, se abrieron más preguntas que respuestas. ¿Cómo interactúa un sistema nervioso y uno periférico para lograr un comportamiento? ¿Cuáles son los fenómenos físicos específicos involucrados en la generación del canto en el sistema animal específico que nos interesa estudiar? ¿Está lista la tecnología para procesar en tiempo real señales fisiológicas, y realizar los cálculos necesarios para emular la tarea comportamental deseada? ¿Cómo cuantificar el acuerdo entre las simulaciones y el canto? ¿Son perceptualmente similares? Algunas de estas preguntas se resolvieron, otras quedan planteadas, así como el alcance de nuestra tecnología para avanzar sobre prótesis humana.

La generación del canto de las aves requiere la coordinación de instrucciones para varios sistemas motores. Principalmente, el sistema respiratorio y el sistema a cargo del control de los músculos siringeos. Sin embargo, no todas las propiedades acústicas

en el canto pueden ser fácilmente identificadas en las instrucciones de control. Parte de las propiedades acústicas emergen del hecho que los sistemas motores conducen un órgano vocal con un alto comportamiento no-lineal. Más aún, en un tema que es recurrente en la biología, distintas especies emplean estrategias diferentes para realizar una tarea. El modo en que los músculos siríngeos opera con el fin de controlar algunas propiedades acústicas en el canto de algunas especies, como el cardenal (Goller and Larsen, 1997), o el brown thrasher (Goller and Suthers, 1996), no es el mismo que con el que opera para controlar la frecuencia en el caso de una de las especies más estudiadas: el diamante mandarín.

La elección de una especie al momento de encarar un estudio en biología está guiada por una diversidad de factores, algunos sólo muy vagamente relacionados con la pregunta en cuestión. Por ejemplo, el diamante mandarín es fácil de criar en el laboratorio, se reproduce durante todo el año, y requiere de relativamente poco tiempo de exposición a un tutor para aprender. Estas características lo han hecho una especie popular y una gran cantidad de información está disponible sobre su comportamiento, su anatomía y su fisiología. Por estos motivos, es difícil resistirse a la tentación de elegirla a la hora de decidirse por una especie. Sin embargo, en esta tesis tuvimos que dilucidar mecanismos singulares empleados por el Diamante Mandarín a la hora de producir el canto. Esta especie, a diferencia de otras estudiadas en la literatura previamente, alterna sílabas tonales con sonidos de altísimo contenido espectral, y para nuestra sorpresa, los mecanismos dinámicos involucrados en la generación de estos últimos, se apartan de aquellos previamente estudiados en la literatura (Gardner et al., 2001; Mindlin and Laje, 2005). De este modo, estuve claramente expuesto a la tremenda diversidad de soluciones que distintas especies encuentran a la hora de generar comportamiento. En este marco, en esta tesis, se muestra que en el canto del Diamante Mandarín, existe una relación entre el contenido espectral y la frecuencia fundamental. Nuestra hipótesis es que esta relación emerge de la dinámica propia del órgano vocal. Para testear dicha hipótesis, se extendió el modelo de la siringe considerando componentes no-lineales en las fuerzas descriptas.

En este punto, la dinámica no lineal se nos presentó como una herramienta insustituible para analizar y generalizar nuestros resultados. En la dinámica, uno aprende a clasificar soluciones independientemente de los detalles de las ecuaciones que las mismas satisfacen. Existen, por ejemplo, para sistemas de baja dimensionalidad, un conjunto finito de modos de encender oscilaciones. Como los comportamientos temporales de los labios responsables de modular el flujo de aire a la hora de vocalizar son regulares, es una hipótesis sensata asumir que la dimensionalidad de las ecuaciones que rigen su dinámica es baja. Asumiendo esto, uno puede analizar si los comportamientos observados son compatibles con alguno de los pocos modos en que un sistema de baja dimensión puede comenzar a oscilar, e intentar recuperar, a partir de los datos, pistas que permitan elegir entre uno u otro. A su vez, una vez identificado el proceso dinámico en juego, es posible investigar si alguno de los mecanismos físicos plausibles detrás del fenómeno en estudio es compatible con la dinámica identificada. De este modo, avanzar en el modelado sabiendo que uno podrá

---

reproducir las propiedades cualitativas de interés. En efecto, procediendo de acuerdo con esta estrategia, hemos realizado una modificación de los modelos previamente presentados en la literatura, para ser capaces de elaborar un modelo físico capaz de dar cuenta de los sonidos producidos por la especie en estudio. Más específicamente, hemos realizado una modificación del modelo de modo que el mismo presente un nuevo régimen dinámico, la aparición de oscilaciones mediante la bifurcación SNILC, capaz de dar cuenta de la relación observada entre frecuencia fundamental y contenido espectral en las vocalizaciones de un Diamante Mandarín. A su vez, este resultado tiene una lectura interesante más allá de lo específico: ilustra en forma contundente que no toda propiedad de un comportamiento es controlada en forma independiente por el sistema nervioso. En nuestro ejemplo, contenido espectral y frecuencia fundamental no son propiedades a sintonizar mediante ajustes específicos de redes neuronales específicas, sino que se presentan vinculadas por propiedades biomecánicas del aparato periférico.

Desde un punto de vista técnico, las oscilaciones encontradas en esta bifurcación (SNILC) nacen al superar la presión un valor umbral, con una amplitud definida, frecuencia cero y forma pulsátil. El contenido espectral del sonido generado, depende de la presión en los sacos aéreos y de la tensión de los labios. A medida que éstos aumentan, el contenido espectral disminuye. Este efecto es más notorio en las sílabas de baja frecuencia. Para los sonidos de alta frecuencia, el uso de control motor adicional prácticamente no afecta el contenido espectral. Por otro lado, las oscilaciones nacidas en una bifurcación de Hopf no muestran una dependencia similar entre contenido espectral y frecuencia. Por ello, los modelos previamente presentados en la literatura no lograron reproducir las vocalizaciones de especies con un canto espectralmente más complejo ([Gardner et al., 2001](#)). Este nuevo modelo permite sintetizar sílabas de baja frecuencia y alto contenido espectral, las cuales habían sido previamente reportadas en trabajos experimentales ([Jensen et al., 2007](#)).

La descripción de este régimen de oscilaciones nacidas en una bifurcación SNILC en el marco de producción vocal, ha sido el primero en la literatura. El modelo presentado permite el estudio de la producción del canto en especies que, como el Diamante Mandarín, presentan un canto más complejo que el que desarrollan las estudiadas hasta el momento. Considerando que el mecanismo de fonación es análogo en los humanos, resultará interesante también analizar la relevancia de este régimen en la fisio-patología de sus vocalizaciones.

Más allá de las ventajas de la dinámica a la hora de identificar un comportamiento, la misma es capaz de generar una herramienta invaluable: las formas normales. Estas ecuaciones son las más sencillas capaces de presentar una bifurcación específica. En un programa como el nuestro, que apunta al aprovechamiento máximo de una tecnología consistente en integrar ecuaciones en tiempo real, la posibilidad de contar con ecuaciones con un número minimal de términos, pero capaces de conservar propiedades dinámicas específicas, se mostró como una herramienta teórica de gran valor.

Partiendo de la hipótesis que las características acústicas del sonido sintetizado por el modelo emergen de las bifurcaciones presentes en el mismo, se desarrolló

un modelo minimal que presenta la misma dinámica. Utilizando este modelo, se construyó un dispositivo biomimético que transforma en tiempo real instrucciones fisiológicas (actividad en los músculos siringeos) en canción. Dicho dispositivo, diseñado en una placa DSP, lee los datos fisiológicos de entrada, integra el modelo propuesto y genera la salida sonora considerando el filtrado en el tracto superior.

El éxito en la síntesis en tiempo real de canto con características espectrales complejas, como el producido por el Diamante Mandarin, construye confianza en la siguientes hipótesis adoptadas: Las características acústicas dependen de las bifurcaciones presentes en el modelo. La integración de un modelo minimal es suficiente para replicar dichas características. La tecnología DSP utilizada cumple los requisitos para la correcta integración del modelo en tiempo real. Las variables fisiológicas identificadas pueden comandar exitosamente el modelo de modo de sintetizar canto en tiempo real.

Si bien esta tesis muestra que algunos aspectos importantes de la física involucrada en la generación del canto están comprendidos, muchos aspectos quedan por dilucidar. Hemos podido realizar una integración exitosa, asumiendo que existen dos regímenes distintos en juego para generar vocalizaciones de alta y baja frecuencia. Sin embargo, no hemos podido identificar, en ninguna de las variables que hemos medido, cuál es el instrumento fisiológico que opera para transitar de un modo al otro. Ninguna de las actividades musculares registradas pudo asociarse al cambio de régimen; ninguna manipulación intentada pudo inducirla, o eliminarla. Desde el punto de vista estrictamente físico, el canto del Diamante Mandarin sigue presentando interrogantes que no hemos agotado.

Desde un punto de vista estrictamente técnico, en esta tesis mostramos que la tecnología DSP está lista para cumplir con el objetivo propuesto: integrar en tiempo real ecuaciones inspiradas en la biomecánica, entre los tiempos requeridos para la toma de datos fisiológicos. Esto abre una nueva perspectiva para el empleo de tecnología DSP en problemas biológicos.

En el presente, la tecnología DSP está siendo implementada en una gran variedad de problemas de inspiración biológica. Es probable que esta tecnología se convierta en una solución estándar en el desarrollo de aplicaciones biomiméticas. En particular, las aplicaciones BMI (*Brain Machine Interfaces*) funcionan típicamente del siguiente modo: leen variables experimentales y luego realizan análisis estadísticos de modo de obtener los parámetros necesarios para controlar dispositivos periféricos. Dicho enfoque se torna computacionalmente muy demandante. En esta tesis, la solución biomimética implementada en el DSP se basa en la integración de las ecuaciones que modelan la física detrás del problema. Se propone que esta estrategia reduce severamente la demanda computacional, por ende mejorando la capacidad tecnológica actual para aplicaciones biomiméticas.

Con el objetivo de estudiar la integración sensorimotora, el desarrollo de experimentos de feedback auditivo alterado se ha transformado en un objetivo importante para la comunidad de canto de aves. Experiencias previas han tenido varios inconvenientes por tener delays muy largos entre la entrada y la salida del sistema (Leonardo and Konishi, 1999) o dificultades con la flexibilidad en su implementación técnica (Zysman et al., 2005). El dispositivo presentado en esta tesis presenta ventajas respecto enfoques anteriores. Primero resuelve técnicamente los problemas de delay y flexibilidad. Pero aún más importante, su principal ventaja es que intrínsecamente considera la física del problema, por lo que cualquier alteración que se introduzca en el mismo tiene implicancias fisiológicas. Por ello, este dispositivo abre la posibilidad de desarrollar experimentos de feedback auditivo precisos y altamente controlados.

Aún más, esta prótesis vocal exitosamente controlada en tiempo real con señales fisiológicas, puede establecer las bases para el desarrollo de dispositivos similares en humanos. En ese sentido, el uso de la tecnología de DSP y la utilización de modelos físicos, permite fácilmente incorporar mejoras incrementales en calidad del sonido sintetizado. Dichas mejoras se incorporaran al introducir en el modelo progresos en el conocimiento, tanto de la física del proceso como de la fisiología detrás del problema.

El modelado en biología es un desafío permanente. Existe una mínima complejidad por debajo de la cual, la vida, sencillamente no ocurre. Por este motivo, la construcción de dispositivos biomiméticos cumple una doble funcionalidad: nos permite imaginar soluciones aplicables a la salud, así como también comprender si las hipótesis simplificadoras implícitas en cualquier modelo, son un mínimo aceptable desde un punto de vista conceptual. Esfuerzos como los presentados en esta tesis muestran, sin embargo, que el avance hacia una visión más cuantitativa de las ciencias biológicas es posible y probablemente, sea uno de los pilares de las investigaciones futuras en este campo del conocimiento.





# El Sonograma

---

El sonograma (o espectrograma) es la representación en tiempo y frecuencia del sonido. El eje X representa al tiempo y el eje Y la frecuencia. Los sonidos de baja frecuencia están representados cerca de la base y los de alta frecuencia en el sector superior. En este gráfico espectro-temporal del sonido, la información de la amplitud se indica mediante la intensidad en escala de grises, trazas más oscuras indican amplitudes mayores. La escala de grises puede ser reemplazada indistintamente por una escala de colores. El sonograma es una herramienta ideal para representar el canto de las aves. Permite fácilmente interpretar la estructura temporal de componentes acústicos. Antes de la introducción de esta técnica, la interpretación se realizaba utilizando onomatopeyas o partituras musicales.

La construcción del sonograma se realiza del siguiente modo: se parte la serie temporal en ventanas equiespaciadas; en cada de las ventanas se calcula la transformada de Fourier; se calcula el módulo al cuadrado de la transformada y se lo coloca, en  $t$  correspondiente del eje X, como una serie de puntos de distinta intensidad en el eje Y. Típicamente, el cálculo se realiza sobre una señal discreta, por lo que el cálculo se realiza utilizando la siguiente expresión:

$$X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}$$

donde  $x[n]$  es la señal a analizar y  $w[n]$  es la función que define la ventana. En este caso  $m$  es discreta y  $\omega$  es continua, pero en la mayoría de las aplicaciones, la transformada se realiza utilizando una FFT discreta. Existen varias opciones para la función de la ventana. En general, se utiliza una función Gaussiana o de Hann. El ancho de la ventana se ajusta en función del sonido. Una ventana más ancha brinda precisión en el espacio de la frecuencia, pero la pierde en el espacio del tiempo. Cuando la ventana es angosta la situación es inversa.

Técnicamente, calculamos el sonograma de la grabaciones utilizando programas de distribución libre o mediante una rutina de MATLAB.

- **Praat** ( <http://www.fon.hum.uva.nl/praat/> )
- **SND** ( <http://ccrma.stanford.edu/software/snd/> )
- **MATLAB** (  $S = spectrogram(x)$  )



# Bibliografía

- Amador A, Goller F, Mindlin G (2008) Frequency modulation during song in a suboscine does not require vocal muscles. *Journal of Neurophysiology* 99:2383. 34
- Amador A, Mindlin G (2008) Beyond harmonic sounds in a simple model for birdsong production. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 18:043123. 35, 36
- Arneodo E, Mindlin G (2009) Source-tract coupling in birdsong production. *Physical Review E* 79:61921. 15, 18
- Basmajian J (1962) Muscles alive. Their functions revealed by electromyography. *Academic Medicine* 37:802. 3
- Blomgren M, Chen Y, Ng M, Gilbert H (1998) Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *The Journal of the Acoustical Society of America* 103:2649. 14
- Boersma P, Weenink D (2001) Praat, a system for doing phonetics by computer. *Glott International* 5:341–345. 57
- Bottjer S, Miesner E, Arnold A (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224:901. 10
- Catchpole C, Slater P, Mann N (2003) *Bird song: biological themes and variations* Cambridge Univ Pr. 4
- Chassaing R (2002) *DSP applications using C and the TMS320C6x DSK* Wiley-Interscience. 54, 62
- Chiel H, Beer R (1997) The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences* 20:553–557. 2, 6, 9
- Clark G (2003) *Cochlear implants: fundamentals and applications* Springer Verlag. 3
- Daley M, Goller F (2004) Tracheal length changes during zebra finch song and their possible role in upper vocal tract filtering. *Journal of neurobiology* 59:319–330. 14, 34
- Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience* 22:567–631. 4
- Elemans C, Mead A, Rome L, Goller F (2008) Superfast vocal muscles control song production in songbirds. *PLoS ONE* 3. 12

- Fee M, Shraiman B, Pesaran B, Mitra P (1998) The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature* 395:67–71. 14
- Fletcher N, Tarnopolsky A (1999) Acoustics of the avian vocal tract. *The Journal of the Acoustical Society of America* 105:35. 13
- Gardner T, Cecchi G, Magnasco M, Laje R, Mindlin G (2001) Simple motor gestures for birdsongs. *Physical Review Letters* 87:208101. 11, 16, 23, 35, 66, 67
- Gardner T, Naef F, Nottebohm F (2005) Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308:1046. 21
- Goldman S, Nottebohm F (1983) Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. *Proceedings of the National Academy of Sciences USA* 80:2390–2394. 5
- Goller F, Cooper B (2004) Peripheral motor dynamics of song production in the zebra finch. *ANNALS-NEW YORK ACADEMY OF SCIENCES* pp. 130–152. 13, 27, 56, 60
- Goller F, Daley M (2001) Novel motor gestures for phonation during inspiration enhance the acoustic complexity of birdsong. In *Proceedings. Biological sciences/The Royal Society*, Vol. 268, p. 2301. The Royal Society. 14
- Goller F, Larsen ON (1997) A new mechanism of sound generation in songbirds. *Proceedings of the National Academy of Sciences* 94:14787–14791. 12, 66
- Goller F, Mallinckrodt M, Torti S (2004) Beak gape dynamics during song in the zebra finch. *Developmental Neurobiology* 59:289–303. 15, 24
- Goller F, Suthers R (1996) Role of syringeal muscles in controlling the phonology of bird song. *Journal of neurophysiology* 76:287. 14, 66
- Golubitsky M, Stewart I, Buono P, Collins J (1999) Symmetry in locomotor central pattern generators and animal gaits. *Nature* 401:693–695. 1
- Guckenheimer J, Holmes P (1990) *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields* Springer. 35, 42
- Hartley R, Suthers R (1989) Airflow and pressure during canary song: direct evidence for mini-breaths. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* 165:15–26. 13
- Hofbauer J, Sigmund K (1988) The theory of evolution and dynamical systems. Mathematical aspects of selection. *CAMBRIDGE UNIVERSITY PRESS, NEW YORK, NY(USA)*. 1988. . 1
- Horowitz P, Hill W (1989) *The art of electronics* Cambridge Univ Pr. 49, 52

- Izhikevich E (2006) *Dynamical systems in neuroscience: The geometry of excitability and bursting* The MIT press. 1, 29
- Jarvis E, Ribeiro S, Da Silva M, Ventura D, Vielliard J, Mello C (2000) Behaviourally driven gene expression reveals song nuclei in hummingbird brain. *Nature* 406:628. 9
- Jensen K, Cooper B, Larsen O, Goller F (2007) Songbirds use pulse tone register in two voices to generate low-frequency sound. *Proceedings of the Royal Society B* 274:2703. 14, 27, 56, 67
- Junge D Nerve and muscle excitation. *Cambridge, MA* . 3
- Keener J, Sneyd J (2008) *Mathematical Physiology: Cellular Physiology* Springer. 1
- Kempermann G, Wiskott L, Gage F (2004) Functional significance of adult neurogenesis. *Current opinion in neurobiology* 14:186–191. 5
- King A (1989) Functional anatomy of the syrinx. *Form and function in birds* 4:105–192. 10
- Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift für Tierpsychologie* 22:770–783. 4
- Laje R, Gardner T, Mindlin GB (2002) Neuromuscular control of vocalizations in birdsong: a model. *Physical Review E* 65:51921. 17, 23, 33
- Laje R, Mindlin G (2005) Modeling source-source and source-filter acoustic interaction in birdsong. *Physical Review E* 72:36218. 18
- Larsen ON, Goller F (1999) Role of syringeal vibrations in bird vocalizations. *Proc. R. Soc. Lond. B* 266:1609–1615. 11, 12, 58
- Larsen O, Goller F (2002) Direct observation of syringeal muscle function in songbirds and a parrot. *Journal of Experimental Biology* 205:25. 12
- Lebedev M, Nicolelis M (2006) Brain-machine interfaces: past, present and future. *TRENDS in Neurosciences* 29:536–546. 3
- Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* 399:466–469. 69
- Margoliash D (1997) Functional organization of forebrain pathways for song production and perception. *Journal of neurobiology* 33:671–693. 5
- Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology* 71:1–25. 4

- May R, Oster G (1976) Bifurcations and dynamic complexity in simple ecological models. *The American Naturalist* 110:573. 1
- McCasland J, Konishi M (1981) Interaction between auditory and motor activities in an avian song control nucleus. *Proceedings of the National Academy of Sciences* 78:7815. 9
- Millán J, Renkens F, Mouriño J, Gerstner W (2004) Noninvasive brain-actuated control of a mobile robot by human EEG. *IEEE Transactions on Biomedical Engineering* 51:1026–1033. 3
- Mindlin GB, Solari HG, Natiello MA, Gilmore R, Hou XJ (1991) Topological analysis of chaotic time series data from the belousov zhabotinskii reaction. *Journal of Nonlinear Science* 1:147–153. 29
- Mindlin G, Gardner T, Goller F, Suthers R (2003) Experimental support for a model of birdsong production. *Physical Review E* 68:41908. 21, 23, 35, 58
- Mindlin G, Laje R (2005) *The physics of birdsong* Springer Verlag. 12, 19, 22, 29, 66
- Nottebohm F, Stokes T, Leonard C (1976) Central control of song in the canary, *Serinus canarius*. *J Comp Neurol* 165:457–486. 9
- Nowicki S (1987) Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere . 18
- Paja C, Romero A, Giral R (2007) Evaluation of Fixed-Step Differential Equations Solution Methods for Fuel Cell Real-Time Simulation In *Clean Electrical Power, 2007. ICCEP'07. International Conference on*, pp. 480–487. 62
- Perelson A (2002) Modelling viral and immune system dynamics. *Nature Reviews Immunology* 2:28–36. 1
- Poole J, Tyack P, Stoeger-Horwath A, Watwood S (2005) Elephants are capable of vocal learning. *Nature* 434:455–456. 4
- Press W, Teukolsky S, Vetterling W, Flannery B (1992) *Numerical recipes in C* Cambridge Univ. Press Cambridge MA, USA.: 55, 62
- Randall D, Burggren W, French K, Eckert R (2001) *Eckert animal physiology: mechanisms and adaptations*, chapter 10 WH Freeman. 11
- Riede T, Suthers R, Fletcher N, Blevins W (2006) Songbirds tune their vocal tract to the fundamental frequency of their song. *Proceedings of the National Academy of Sciences* 103:5543. 18
- Sanz J, Arneodo EM, Mindlin GB (2009) Source-tract interactions in zebra finch beak dynamics. 15

- Schlick T (2002) *Molecular modeling and simulation* Springer New York. 1
- Sitt J, Aliaga J (2007) Versatile biologically inspired electronic neuron. *Physical Review E* 76:51919. 62
- Sitt J, Amador A, Goller F, Mindlin G (2008) Dynamical origin of spectrally rich vocalizations in birdsong. *Physical Review E* 78:11905. 41
- Smith S (1999) The scientist and engineer's guide to digital signal processing . 54
- Smolen P, Baxter D, Byrne J (2000) Mathematical modeling of gene networks. *Neuron* 26:567–580. 1
- Solari H, Natiello M, Mindlin G (1996) *Nonlinear dynamics: a two-way trip from physics to math* Taylor & Francis. 42
- Spiro J, Dalva M, Mooney R (1999) Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *Journal of neurophysiology* 81:3007. 5, 11
- Strogatz S (2001) *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering* Perseus Books. 18, 29, 44
- Suthers R (2001) Peripheral vocal mechanisms in birds: are songbirds special? *Netherlands Journal of Zoology* 51:217–242. 10, 12
- Suthers R, Goller F, Pytte C (1999) The neuromuscular control of birdsong. *Philosophical Transactions of the Royal Society B: Biological Sciences* 354:927. 12
- Tchernichovski O, Mitra P, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291:2564. 14
- Thorpe W (1958) The learning of song patterns by birds, with especial reference to the song of the Chaffinch *Fringilla coelebs*. *Ibis* 100:535–570. 5
- Titze I Principles of voice production. 1994. 11
- Titze I (1988) The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am* 83:1536–1552. 16, 39
- Tomlin C, Axelrod J (2007) Biology by numbers: mathematical modelling in developmental biology. *Nature reviews genetics* 8:331–340. 1
- Vicario D (1991) Contributions of syringeal muscles to respiration and vocalization in the zebra finch. *J. Neurobiology* 22:63–73. 14, 34, 37, 52
- Webb B, Consi T (2001) *Biorobotics: methods and applications* AAAI Press/MIT Press. 2



- Wessberg J, Stambaugh C, Kralik J, Beck P, Laubach M, Chapin J, Kim J, Biggs S, Srinivasan M, Nicolelis M (2000) Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* 408:361–365. 3
- White SA (2009) Genes and vocal learning. *Brain and Language* In Press, Corrected Proof:–. 6
- Williams H (2004) Birdsong and singing behavior. *Ann. N.Y. Acad. Sci.* 1016:1–30. 14
- Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. *Science* 273:1871–1875. 10
- Zrenner E (2002) Will retinal implants restore vision? *Science* 295:1022–1025. 3
- Zysman D, Méndez J, Pando B, Aliaga J, Goller F, Mindlin G (2005) Synthesizing bird song. *Physical Review E* 72:51926. 49, 62, 63, 69