

UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Física

La Física del Canto de Aves

por Rodrigo Laje

Director: Dr. Gabriel B. Mindlin
Lugar de trabajo: Departamento de Física, FCEyN, UBA

Trabajo de Tesis para optar por el título de
Doctor de la Universidad de Buenos Aires

Agosto 2005

Resumen

Esta tesis está dedicada al análisis y modelado de los mecanismos físicos de la fonación en el órgano vocal de las aves canoras, la *siringe*.

Algunas aves son popularmente conocidas por su capacidad de imitar sonidos. Se trata de todas las especies del suborden Oscinos (aves canoras), más los loros y los colibríes. Más allá del hecho curioso, la gran cantidad de similitudes entre la adquisición del habla en humanos y el aprendizaje del canto convierte a estas aves en un modelo invaluable para estudiar los mecanismos neuronales del habla en particular, y de aprendizaje y control motor de un comportamiento complejo en general. Entender cómo el cerebro del ave organiza las acciones motoras del canto es una motivación central en dicho programa.

El vínculo entre actividad neuronal y comportamiento observado lo establece la siringe, el aparato vocal de las aves. La siringe en Oscinos produce sonido por vibración de los *labia* (tejidos móviles análogos a las cuerdas vocales humanas); es un dispositivo mecánico que muestra una riquísima variedad de soluciones oscilatorias. La comprensión de los mecanismos físicos de la fonación es de fundamental importancia para establecer el mapa de control motor y evaluar posibles restricciones a nivel periférico.

En esta tesis presentamos un modelo matemático de fonación en aves canoras cuyos parámetros tienen estrecha relación con parámetros fisiológicos relevantes, como actividad muscular y presión del saco aéreo. Este modelo está basado en un detallado conocimiento de la función de diferentes estructuras en la siringe, gracias a recientes mediciones electrofisiológicas *in vivo*. Proponemos además que una gran variedad de vocalizaciones pueden interpretarse en términos de dos gestos motores cíclicos básicos: la presión del saco aéreo y la tensión de los labia. El parámetro clave para generar sílabas diversas es el desfase entre los gestos motores propuestos.

El incipiente estudio de sonidos complejos en aves (frecuencias subarmónicas, espectros anarmónicos y saltos de frecuencia, por ejemplo) dió por tierra con la idea tácitamente aceptada de que hay una relación directa entre las instrucciones neuronales y las propiedades acústicas del sonido emitido. Los orígenes físicos de la complejidad, sin embargo, distan de ser entendidos completamente. Aquí presentamos un modelo que permite estudiar el acople acústico entre fuente sonora y tracto vocal, así como la interacción acústica entre las dos fuentes sonoras presentes en la siringe. Mostramos que el acople acústico y la realimentación acústica retrasada son posibles causas de sonidos complejos en un sonograma, y proponemos experimentos sencillos para evaluar los orígenes de la complejidad en el canto de algunas aves.

Por último, la idea de los gestos motores nos da un marco para analizar el sorprendente canto a dúo del hornero. Más allá de la física del aparato fonador, la hipótesis de los gestos motores nos permite intuir comportamientos no lineales en los centros neuronales que controlan el canto. Este resultado es una de las pocas evidencias de sincronización de ritmos corporales en animales intactos.

Palabras clave: canto de aves - física de la siringe - control motor

Title: The physics of birdsong

Author: Rodrigo Laje

Abstract

This thesis is dedicated to the analysis and modeling of the physical mechanisms of phonation in the avian vocal organ, the *syrinx*.

Some birds are known due to their ability to imitate sounds. We are referring to songbirds (all species in suborder Oscines), together with parrots and hummingbirds. Beyond the simple curiosity, the large number of similarities between speech acquisition in humans and song learning in these birds makes them an invaluable model to study the neural mechanisms underlying speech and, more generally, learning and motor control of a complex behaviour. Understanding the mechanisms by which the bird brain organizes the song motor gestures is a central motivation in such a program.

The link between neural activity and observed behaviour is established by the syrinx, the avian vocal organ. The syrinx in Oscines generates sound through vibration of the *labia* (mobile tissues analogous to the human vocal cords); it is a mechanical device displaying a rich variety of oscillatory solutions. The comprehension of the physical mechanisms playing at phonation is of fundamental importance in order to build the motor control map and evaluate possible peripheral constraints.

In this thesis we present a mathematical model for the phonation in songbirds whose parameters are closely related to relevant physiological parameters such as muscular activity and air sac pressure. This model is based upon a detailed knowledge of the function of different structures in the syrinx, due to recent *in vivo* electrophysiological measurements. We propose in addition that a large number of different vocalizations can be interpreted in terms of two cyclic, basic motor gestures: air sac pressure and labia tension. The key parameter needed to generate a diversity of syllables is the phase difference between the proposed motor gestures.

The incipient study of complex sounds in birds (subharmonic frequencies, anharmonic spectra and frequency jumps, for instance) abolished the unexpressed belief of a direct relationship between neural instructions and acoustic properties of the emitted sound. The physical origins of complexity, however, remain obscure. Here we present a model that allows us to study the acoustic coupling between sound source and vocal tract, as well as the acoustic interaction between the two sound sources in the syrinx. We show that acoustic coupling and delayed acoustic feedback are possible origins of complex sounds in a sonogram, and propose simple experiments to evaluate the source of complexity in the song of some birds.

Last, the motor gestures hypothesis gives us a framework to study the surprising song of the duetting bird hornero. Beyond the physics of the vocal organ, we found a nonlinear behaviour that allows us to have an insight into the neural circuits driving the song. This result is one of the few evidences of synchronization of bodily rhythms in intact animals.

Keywords: birdsong - physics of the syrinx - motor control

Prefacio

La gente cree que nos divertimos demasiado con el canto de los pájaros. Tengo que decir que... es cierto. Bueno, yo por lo menos me divierto, y estoy seguro de que mis compañeros de grupo también. Es que uno no tiene frecuentemente la oportunidad de aunar el trabajo que le gusta con un tema apasionante y tan al alcance de la mano. Después de todo, los horneros siempre estuvieron ahí, en la ventana, cantando a dúo hasta que Gabo dijo: “Oíste? Parece como si intercalaran una sílaba cada tres...”

Desde la física de la siringa hasta el control motor, desde los sonidos complejos hasta los mecanismos cerebrales de aprendizaje y realimentación auditiva, desde los experimentos de comportamiento hasta los de neurofisiología, el canto de aves es un mundo de cosas para entender. Uno se siente como un cachorro en el cuarto desordenado de su pequeño dueño: tantas cosas interesantes para olfatear y morder.

(Hasta acá, un lecho de rosas. Tendría que hablar del largo esfuerzo, del trabajo duro y sostenido, de los intentos infructuosos, de los desacuerdos, de las metidas de pata, de las horas de computadora. Y de los bajos sueldos, y los platos que de todas formas tenemos que lavar. Lo mejor es que podría seguir nombrando “desventajas” y a pesar de ello conservar una sonrisa de satisfacción.)

Esta tesis, más que nada, es el comienzo.

Índice General

Resumen	iii
Abstract	v
Prefacio	vii
1 Introducción	1
1.1 El canto de las aves	1
1.1.1 Una capacidad poco común	1
1.1.2 Modelo de aprendizaje vocal	2
1.1.3 Entendiendo el aparato fonador	3
1.1.4 Un acercamiento novedoso	3
1.2 La física de la fonación en aves	4
1.2.1 El fin de las suposiciones	4
1.2.2 Modelos de fonación	5
1.2.3 Gestos motores	5
1.2.4 Otros orígenes de complejidad	5
1.2.5 Más allá de la física	6
1.3 Esta tesis	6
2 La siringe	9
2.1 Forma y función	9
2.1.1 Morfología de la siringe	9
2.1.2 Mecanismo general de producción de sonido	10
2.2 Oscinos	12
2.2.1 El rol de los músculos	13
2.2.2 Lateralización	14
2.2.3 Aves que aprenden y musculatura intrínseca	17
2.3 No-Oscinos	17
3 Fonación	19
3.1 Un poco de historia	19
3.1.1 Modelos vibrantes	19
3.1.2 La hipótesis del silbido	20
3.1.3 Modelos de fonación en humanos	21
3.2 Experimentos definitivos	22
3.2.1 Experimentos con heliox	22
3.2.2 Endoscopía y lesiones	23

3.3	Oscilaciones en la siringe	23
3.3.1	Por qué modelar la siringe	23
3.3.2	Oscilaciones autosostenidas	24
3.3.3	Modelo flameante de Titze	25
3.3.4	Modelo extendido	27
3.4	Respaldo experimental	32
3.4.1	Una predicción	32
3.4.2	Un experimento	33
3.5	El tracto vocal	33
3.5.1	El tracto vocal como filtro	34
3.5.2	Cómo filtrar	36
3.5.3	Cuando el tracto vocal es más que un filtro	37
4	Gestos motores para el canto	39
4.1	La estructura del canto	39
4.1.1	Sílabas	39
4.1.2	Patrones de respiración	40
4.2	Gestos para una sílaba	42
4.2.1	Construyendo una sílaba	42
4.2.2	Sílabas y diferencias de fase	44
4.2.3	El control activo del flujo de aire	44
4.2.4	Sílabas con diferente patrón respiratorio	48
4.3	De la siringe al cerebro	48
4.3.1	El cerebro de los Oscinos	48
4.3.2	Sustrato neuronal para la diferencia de fase	51
5	Sonidos complejos	53
5.1	Sonidos complejos	53
5.1.1	Instrucciones vs. mecánica	54
5.1.2	Subarmónicos	54
5.1.3	Paradigma de estructura intrínseca	56
5.1.4	Realimentación retrasada	56
5.2	Interacción acústica fuente-tracto	57
5.2.1	Acople acústico entre fuente y tracto	58
5.2.2	Modelo para una fuente acoplada al tracto	61
5.2.3	Predicciones	63
5.3	Interacción acústica fuente-fuente	66
5.3.1	Fuentes interactuantes	66
5.3.2	Modelo para dos fuentes acopladas al tracto	69
5.3.3	Simulaciones	70
6	Ritmos complejos	73
6.1	Los horneros	73
6.1.1	El dúo de horneros	74
6.1.2	Comportamiento conjunto	75
6.2	Horneros y osciladores forzados	77
6.2.1	Osciladores lineales vs. no lineales	77
6.2.2	Lenguas de Arnold y la escalera del diablo	78
6.2.3	Dúos de prueba	79
6.2.4	El oscilador en Horneros	80

<i>ÍNDICE GENERAL</i>	xi
7 Conclusiones	83
A Detalles	87
A.1 Acople acústico entre fuente y tracto	87
A.2 Modelo de una fuente	89
A.2.1 Forma explícita	89
A.2.2 Aproximación de bajo acople	90
A.3 Modelo de dos fuentes	91
A.3.1 Forma explícita	91
A.3.2 Aproximación de bajo acople	92
Bibliografía	95
Agradecimientos	103

Capítulo 1

Introducción

*Lo que tiene que hacer el auditor
es guardar un silencio religioso
porque tu canto sabe adónde va
perfectamente.*

Nicanor Parra, Defensa de Violeta Parra (1969)

1.1 El canto de las aves

1.1.1 Una capacidad poco común

La capacidad vocal imitativa de algunas aves es popularmente conocida, como la de los loros y los mirlos, por ejemplo. No es tan conocido, sin embargo, que muy pocas otras especies en el reino animal comparten esta sobresaliente capacidad. Con excepción de los humanos, ningún primate aprende sus vocalizaciones; entre el resto de los mamíferos, sólo cetáceos y murciélagos son capaces de aprender de sus tutores [Brainard and Doupe 2002].¹ En contraste, casi la mitad de las especies de aves que existen aprende a vocalizar. Se trata de los varios miles de especies que componen el suborden Oscinos (llamadas “aves canoras”; por ejemplo canarios, zorzales, chingolos, mirlos, etc.), más los loros (orden Psittaciformes) y los colibríes (familia Trochilidae).

El aprendizaje del canto en aves tiene muchos puntos en común con la adquisición del habla por parte de los humanos. Como nosotros, las aves deben escuchar las vocalizaciones de sus tutores durante un período de memorización llamado “sensorial”, y deben ser capaces de oírse durante un período posterior de práctica llamado “sensorimotor”. Durante la fase sensorial, el ave forma una representación interna de lo que oye de sus tutores (el *templado*), mientras que en la fase sensorimotora comparan sus propias vocalizaciones con el *templado*

¹Evidencia reciente mostraría que los elefantes también tienen la capacidad de aprender vocalizaciones [Poole et al. 2005].

adquirido y corrigen errores hasta lograr una copia estereotipada del canto de su especie [Brainard and Doupe 2002]. Los primeros estudios sobre aprendizaje vocal en aves comenzaron con los trabajos de Thorpe y Marler [Thorpe 1961, Marler 1970], quienes demostraron que aves canoras criadas en aislamiento acústico producían cantos anormales (diferentes a los de su especie). Si el ave era criada por adultos conespecíficos (*i.e.* de la misma especie, no necesariamente sus progenitores), producía cantos muy similares a los de sus tutores.

1.1.2 Modelo de aprendizaje vocal

Estas similitudes fundamentales entre el aprendizaje del canto en aves y la adquisición del habla en humanos hacen del canto de pájaros uno de los pocos modelos para el estudio de este mecanismo en humanos [Doupe and Kuhl 1999]. Además hay otras características, que nombraremos a continuación, que se unen a las ya descritas para hacer del canto de aves un modelo invaluable en el estudio de los mecanismos de aprendizaje de un comportamiento complejo.

La primera y más evidente de estas características es la facilidad para obtener y analizar el comportamiento (el canto). Una simple grabación sonora y un programa estándar de análisis de audio son suficientes para comenzar. El canto es ya, en general, un comportamiento relativamente simple y repetitivo. Otras características, no tan evidentes, son acaso más importantes. Los centros neuronales que gobiernan el comportamiento, tanto los de control motor como los relacionados con el aprendizaje y la realimentación auditiva, son pocos en número y están altamente especializados [Farries 2001, Spiro et al. 1999]. Además, los músculos involucrados en la fonación son pocos: la configuración de la delicada estructura vocal en las aves que aprenden está gobernada por apenas 6 pares de músculos, como máximo [Gaunt 1983]. Brenowitz y sus colaboradores [Brenowitz et al. 1997] ofrecen un resumen de las ventajas del canto en aves como modelo para identificar mecanismos neuronales que subyacen al comportamiento; entre otras,

- El canto es un comportamiento repetible y fácilmente registrable.
- La configuración del aparato vocal está gobernada por un número pequeño de músculos.
- El canto es un comportamiento aprendido que está controlado por un pequeño número de centros neuronales discretos.
- El canto es el resultado de programas motores estereotipados, con algún nivel de organización jerárquica de los núcleos motores y premotores.
- El desarrollo del canto se da en fases bien diferenciadas, y se puede relacionar la ontogenia del canto con el desarrollo de los circuitos neuronales subyacentes.
- El canto adulto es un comportamiento plástico que incluye neurogénesis continua y cambios estacionales.
- Muchos aspectos del canto difieren entre las especies que aprenden, lo que provee oportunidades de estudios comparativos.

Como lo señala el último punto, sin embargo, no todo es similitud. No todas las especies en los grupos mencionados tienen las mismas habilidades vocales.

Los humanos, los más prolíficos vocalmente, pueden aprender un número aparentemente infinito de combinaciones de sonidos. No tan prolíficos, los loros y las calandrias, por ejemplo, pueden producir cientos o miles de combinaciones de sonidos aprendidos. Los menos prolíficos vocalmente, como algunas aves canoras muy estereotipadas y los colibríes, tienen un solo tipo de canto con muy poca variación. Hay también otras diferencias insoslayables, como por ejemplo en las duraciones de los períodos de aprendizaje, o las referidas al control del aparato vocal. El estudio comparativo de especies diversas con características tanto similares como diferentes permitirá entender qué es lo especial en el cerebro de estos animales que pueden imitar sonidos, en particular los humanos [Jarvis 2004, Perkel 2004].

1.1.3 Entendiendo el aparato fonador

El aparato fonador de las aves es la *siringe*, una delicada estructura cartilaginosa ubicada en la unión de los bronquios y la tráquea. La fonación ocurre básicamente cuando el aire expelido por los sacos aéreos pasa a través de la siringe, induciendo la vibración de ciertos tejidos móviles (llamados *labia* en los Oscinos). Dicha vibración modula a su vez el flujo de aire, que ingresa a la tráquea como una onda acústica. Lo que percibimos como sonido es la emisión de la onda acústica que se establece en el tracto vocal.

El canto de aves se caracteriza por la modulación del sonido en un amplio rango de escalas temporales. El comportamiento involucra la traducción de un programa de control motor central a la actividad muscular coordinada de los sistemas motores periféricos: la siringe y los músculos respiratorios [Goller and Cooper 2004, Suthers 1997]. Además, dicho programa de control motor central tuvo que evolucionar para operar dentro de las limitaciones de los aparatos vocal y respiratorio [Zollinger and Suthers 2004]. Entender los mecanismos por los cuales el cerebro organiza este complejo comportamiento es una motivación central en el estudio del aprendizaje y control motor del canto [Fee et al. 1998].

Para complicar un poco el panorama, además del control motor central, un nivel adicional de complejidad en el comportamiento está dado por la dinámica inherente al aparato vocal: está demostrado que variaciones leves en las instrucciones neuronales pueden resultar en una secuencia acústica compleja, debido a la no linealidad del aparato fonador [Fee et al. 1998]. Es necesario entonces un conocimiento profundo del funcionamiento de la siringe y de la manera en que traduce las instrucciones motoras en sonido.

Es de esperar que el conocimiento del funcionamiento de la siringe sea trasladable a todas las aves canoras. La habilidad de un imitador vocal como la calandria para reproducir muchos de los patrones motores y acústicos de cantos heterospecíficos (logra imitaciones muy ajustadas del canto de otras especies utilizando el mismo patrón muscular que el ave modelo) sugiere que las reglas de producción de canto son compartidas por todas las especies [Zollinger and Suthers 2004].

1.1.4 Un acercamiento novedoso

Una variedad de aproximaciones experimentales están dedicadas actualmente a entender cómo se representan en el cerebro del ave la percepción y la producción

del canto, y finalmente cómo la experiencia modifica esas representaciones [McCasland 1987, Brainard and Doupe 2002].² Sin embargo, la enorme cantidad de actividad neuronal involucrada en el canto presenta un desafío creciente para la interpretación. Estas aproximaciones “desde arriba” se enfrentan con una abrumadora cantidad de datos, aún en el caso de las aves donde los centros de control neuronal son más accesibles al análisis dado que están altamente especializados.

Como si fuera poco, está demostrada la cuota adicional de complejidad que el aparato fonador aporta en forma independiente de la que naturalmente conlleva la actividad neuronal, como mencionamos más arriba. Queda claro entonces que es necesaria una comprensión física del funcionamiento de la siringe. A fin de establecer una relación entre actividad neuronal y comportamiento, es esencial conocer en profundidad la mecánica del aparato vocal, el efector de las instrucciones neuronales. Un acercamiento “desde abajo” es novedoso en este campo; además, las herramientas de la Dinámica No Lineal nos permitieron establecer un mojón en este modo de analizar el canto de las aves [Gardner et al. 2001].

1.2 La física de la fonación en aves

Los mecanismos físicos de fonación en aves fueron objeto de especulación durante mucho tiempo, debido a la inaccesibilidad y al pequeño tamaño de la siringe. Básicamente, tres paradigmas de fonación compartieron escenario hasta épocas recientes: el modelo de membranas, el modelo de válvula de presión, y el modelo de silbato. Los tres modelos estaban basados mayormente en evidencia anatómica, pues la función de las diferentes estructuras de la siringe se develó recién con el desarrollo de sofisticadas técnicas de observación *in vivo*. La mayor diferencia entre estos tres modelos radica en que el sonido se produce por la vibración (membranas y válvula de presión) o no (silbato) de alguna parte de la siringe.

1.2.1 El fin de las suposiciones

Las especulaciones sobre la función de las diferentes partes de la siringe, y en particular sobre el mecanismo de fonación de las aves, tuvieron su fin con los trabajos de Goller, Suthers y Larsen. Experimentos de lesiones y endoscopia permitieron determinar que la fonación se debe a la vibración de los labia (en aves canoras [Goller and Larsen 1997b, Larsen and Goller 1999]) o de las LTM (en loros y palomas [Goller and Larsen 1997a, Larsen and Goller 1999]), dándole sustento al modelo de válvula de presión.

En todas las especies de aves estudiadas, la fonación se produce por la vibración de tejidos móviles inducida por el flujo de aire. Bajo ciertas condiciones, el flujo de aire es capaz de entregar energía al tejido móvil; si la energía neta en un ciclo sobrepasa las pérdidas por fricción y choques plásticos, se sostiene la vibración. Dicha vibración a su vez modula al flujo de aire, que ingresa en el tracto vocal como una perturbación acústica. Este mecanismo es análogo al de fonación en humanos.

²Recientemente fue anunciada la creación de un ambicioso proyecto integrador sobre el cerebro del ave que involucra análisis a nivel molecular, anatómico, electrofisiológico y comportamental [Jarvis et al. 2002].

Incluso antes de que el origen del sonido fuera determinado, el rol de los músculos siríngeos en Oscinos había sido determinado con precisión por Goller y Suthers. Con unas asombrosas mediciones *in vivo* [Goller and Suthers 1995, Goller and Suthers 1996a, Goller and Suthers 1996b], dejaron abierto el camino al modelado de la siringe de las aves canoras con base en un conocimiento amplio de la función de diferentes estructuras en ella.

1.2.2 Modelos de fonación

Al contrario que en aves, donde los modelos de fonación tuvieron sustento experimental sólo recientemente, los modelos de fonación en humanos tienen una larga historia bien sustentada en observaciones experimentales. En particular, Titze presentó uno de los modelos más simples [Titze 1988] para tener en cuenta el mecanismo básico de transferencia de energía del flujo de aire a los tejidos móviles. Aunque fue desarrollado sólo para analizar el umbral de las oscilaciones de las cuerdas vocales humanas, este modelo es el punto de partida de nuestro trabajo sobre la fonación en aves canoras.

En esta tesis desarrollamos un modelo matemático de fonación basado en el mencionado modelo de Titze y en los detallados experimentos de Goller, Suthers y Larsen. Los parámetros de nuestro modelo tienen estrecha relación con datos fisiológicos relevantes: actividad muscular en la siringe y presión del saco aéreo [Laje et al. 2002, Suthers and Margoliash 2002]. Este modelo ofrece predicciones concretas sobre actividad muscular a lo largo de la vocalización que fueron validadas experimentalmente [Mindlin et al. 2003].

1.2.3 Gestos motores

Construyendo sobre las preguntas físicas concretas, nuestro trabajo sobre fonación en la siringe nos llevó un paso más arriba hacia el cerebro del ave. Propusimos que una gran variedad de sílabas puede ser interpretada como el resultado de dos gestos motores básicos: la presión del saco aéreo y la tensión de los labios [Gardner et al. 2001]. Estos gestos motores están bien identificados experimentalmente y tienen interpretación directa en términos de los parámetros de nuestro modelo.

Nuestro trabajo propone que una sucesión de sílabas en el canto es el resultado de la aplicación cíclica de estos dos gestos motores básicos. El parámetro clave para generar sílabas diferentes es el desfase entre los gestos motores de presión y tensión. La variedad posible de sílabas aumenta al incorporar un tercer gesto motor: el control del flujo de aire [Laje et al. 2002]. Una implicación de este modelo es que, para relacionar actividad neuronal central con acción motora periférica, sería necesario relacionar actividad central con los gestos motores propuestos, más que con las propiedades acústicas del sonido emitido [Suthers and Margoliash 2002].

1.2.4 Otros orígenes de complejidad

Hasta hace pocos años, la idea de una relación directa entre las propiedades acústicas de la vocalización y las instrucciones neuronales originarias era tácitamente aceptada. Sin embargo, muchas de las propiedades acústicas de la vocalización pueden deberse puramente a la mecánica del aparato fonador sin

tener un correlato de complejidad a nivel de las instrucciones neuronales [Fee et al. 1998]. ¿Cuáles son los posibles orígenes de sonidos complejos a nivel de la siringe?

El paradigma de estructura intrínseca del aparato fonador para explicar sonidos complejos en aves (frecuencias subarmónicas, espectros anarmónicos, saltos de frecuencia, etc.) fue heredado directamente del campo de la voz humana [Fee et al. 1998]. Sin embargo, hay otras causas posibles para la aparición de signaturas de no linealidad en un espectro. Una de las causas posibles es la interacción acústica entre la fuente sonora y el tracto vocal, que puede tener efectos cualitativamente diferentes cuando los rebotes de la onda acústica que se propaga en el tracto vocal vuelven a la fuente sonora e influyen sobre su dinámica, un tiempo después. Estudiamos este efecto por primera vez en un modelo de la voz humana [Laje et al. 2001]. Seguimos esta vertiente y trasladamos el concepto al aparato vocal del ave, analizando el efecto de la realimentación acústica retrasada en la dinámica de la siringe [Laje and Mindlin 2005]. Propusimos experimentos relativamente sencillos que pueden echar luz sobre el origen físico de sonidos complejos en ciertos espectros.

La complejidad en el canto de aves no termina en las soluciones posibles de una fuente sonora. La siringe de los Oscinos es *bilateral*: tiene dos fuentes de sonido que actúan con cierto grado de independencia, y que pueden interactuar [Nowicki and Capranica 1986a]. La interacción deja su huella en el espectro, aunque su origen físico todavía es incierto y podría depender de la especie. Actualmente estamos estudiando el acople acústico entre las dos fuentes sonoras como una causa posible de interacción [Sciamarella et al. 2005].

1.2.5 Más allá de la física

Nuestro trabajo sobre gestos motores en Oscinos nos permitió interpretar un hallazgo llamativo: la organización no trivial de las sílabas de macho y hembra en el dúo de horneros [Laje and Mindlin 2003]. Conocidos por su nido de barro, los horneros son Suboscinos³ que cantan a dúo en una forma que a primera vista parece descoordinada. El dúo, sin embargo, muestra características propias de los osciladores no lineales forzados periódicamente. Al comienzo del dúo, macho y hembra cantan sílabas en proporción 1/1; a medida que el dúo se desarrolla, el macho incrementa su ritmo de producción de sílabas y se encuentran sucesivamente proporciones 2/3, 1/2, 2/5, 1/3, 1/4, 1/5, como las lenguas de comportamiento subarmónico de los osciladores no lineales.

No es mucho lo que se conoce sobre el cerebro de los Suboscinos. Este hallazgo nos sugiere que los gestos motores de la hembra se encuentran acoplados (en principio, a nivel de núcleos medulares) en la forma de un oscilador no lineal o un sistema excitable; además, es evidencia indirecta de una conexión entre los centros auditivos y motores.

1.3 Esta tesis

Esta tesis está dedicada a entender los mecanismos físicos de la fonación en aves canoras. Analizamos el comportamiento del órgano vocal del ave con las

³Son aves que no aprenden. En el árbol filogenético, el suborden Suboscinos es paralelo al suborden Oscinos; ambos subórdenes forman el orden Passeriformes.

herramientas y el espíritu de la Dinámica No Lineal. Desarrollamos modelos de fonación cuyos parámetros tienen estrecha relación con datos fisiológicos, y propusimos que una gran variedad de vocalizaciones son interpretables en términos de un pequeño número de parámetros de control.

Un aspecto central de la Dinámica No Lineal es la búsqueda de modelos paradigmáticos que tengan un comportamiento determinado. Gracias a esta característica, los resultados de nuestra investigación nos permitieron ir más allá de la física del problema para tener una intuición sobre los mecanismos neuronales periféricos de control motor necesarios para generar las vocalizaciones.

Sigue una descripción general sobre la organización de esta tesis.

Capítulo 2: la siringe. Una revisión de lo que se conoce sobre forma y función del aparato vocal de las aves. Describo el rol de los músculos siríngeos en el control de la configuración de la siringe de las aves canoras, y la función de diferentes estructuras en él. Una breve descripción de la siringe de otros grupos (palomas y loros) muestra que puede haber diferencias importantes en el control del aparato vocal.

Capítulo 3: la producción de sonido. En este capítulo presento el modelo de fonación desarrollado en esta tesis. Comienza con una breve reseña sobre la evolución histórica de las hipótesis, los modelos y los experimentos cruciales para explicar el mecanismo de fonación en aves. A continuación sigue el análisis de las condiciones de oscilación de la siringe de Oscinos y el modelo propuesto para describir el comportamiento del aparato vocal, que ofrece acertadas predicciones sobre actividad muscular. Termina con una discusión del papel del tracto vocal en las vocalizaciones de aves.

Capítulo 4: los gestos motores. La idea del canto como resultado de “gestos motores” cíclicos muy simples es desarrollada aquí. Muestro que hay dos parámetros básicos en nuestro modelo (presión del saco aéreo y tensión de los labios) que nos permiten reproducir una gran variedad de vocalizaciones de aves. La diversidad de vocalizaciones se obtiene al variar la diferencia de fase entre dichos gestos motores. Muestro además que la incorporación de un tercer gesto motor (el control activo del flujo de aire) permitiría al ave una mayor variedad de vocalizaciones. Por último, discuto la relación entre los gestos motores propuestos y el sustrato neuronal subyacente en el cerebro de los Oscinos.

Capítulo 5: sonidos complejos. El título se refiere a aquellos sonidos que no podrían ser producidos por un aparato fonador lineal, y que nos dan pistas sobre su comportamiento. Estos sonidos están bastante descriptos en la voz humana, pero no en el canto de aves. Proponemos que la interacción acústica entre las fuentes sonoras de la siringe y el tracto vocal es una posible causa de sonidos complejos. Ofrecemos hipótesis contrastables al respecto que pueden echar un poco de luz sobre los mecanismos no lineales de la siringe.

Capítulo 6: ritmos complejos. En este capítulo describo nuestro trabajo sobre los dúos de horneros. Encontramos que los dúos se organizan de una manera no aparente y no trivial, que es compatible con la forma en que se comportan los osciladores no lineales forzados periódicamente. Este resultado, que nos permite intuir características del cerebro de la hembra

hornero, es una de las pocas evidencias de generación y sincronización de ritmos corporales en animales intactos.

Capítulo 7: conclusiones. Un breve capítulo para discutir los hallazgos y aportes realizados.

Capítulo 2

La siringe

El órgano vocal de las aves, la *siringe*, ha sido objeto de extensa investigación. Los comienzos del análisis científico de la producción vocal en pájaros se remontan al siglo XVIII; los primeros trabajos se dedicaban a la clasificación taxonómica de las aves basada en la morfología de la siringe [Hérissant 1753]. Como veremos a continuación, el aparato vocal de las aves tiene una morfología bastante diversa, aunque el mecanismo básico de fonación es similar en todas las especies estudiadas y es análogo al de las cuerdas vocales humanas.

2.1 Forma y función

2.1.1 Morfología de la siringe

El órgano vocal de las aves, la *siringe*, es una parte diferenciada de las vías aéreas. Está ubicada en el saco aéreo interclavicular, donde los bronquios se unen para formar la tráquea. La siringe está compuesta por estructuras cartilagosas, masas de tejido conectivo, membranas y músculos.

La siringe en Oscinos es, en contraste con el aparato vocal humano, una estructura simétrica bilateral con dos fuentes de sonido: ciertas masas de tejido forman una válvula vibrante de cada lado de la siringe (Figura 2.1a). La parte lateral de cada válvula está formada por el grueso *labium lateral* (LL), una masa de tejido fija al tercer anillo cartilaginoso bronquial. La parte medial es inhomogénea en estructura; está formada por el *labium medial* (ML) en su parte anterior y la delgada *membrana timpaniforme medial* (MTM) en su parte posterior [Fee 2002]. Hay algunas otras especies, además de los Oscinos, que también tienen una siringe doble (Figura 2.1c).

La siringe de palomas, loros y otras especies, en cambio, tiene una única válvula vibrante, formada por la aducción de las *membranas timpaniformes laterales* (LTM; en la Figura 2.1b se muestra la siringe de una paloma). Las LTM son masas de tejido conectivo relativamente gruesas cuya superficie lateral (“externa”) está en contacto con el saco aéreo interclavicular [Goller and Larsen 1997a]. En palomas se encuentran, además, las *membranas timpaniformes mediales* (MTM), que son delgadas y están ubicadas en la pared medial de cada bronquio.

Siringes y taxonomía

La clasificación taxonómica entre Oscinos y no-Oscinos estuvo basada originalmente en la morfología de la siringe [Elemans et al. 2003, Suthers 2001]. Hay tres clases de siringes, diferenciadas por la posición de las membranas fonatorias con respecto a los anillos cartilagosos traqueales y bronquiales (Figura 2.1). La mayoría de las aves poseen una siringe *traqueobronquial* (que es particularmente común en el suborden Oscinos): la característica de este tipo de siringe es que las membranas están sostenidas entre anillos traqueales y bronquiales. La siringe *traqueal*, que sólo involucra anillos traqueales, ocurre en Furnariidae (horneros) y algunos Ciconiidae (cigüeñas). La siringe *bronquial*, sostenida únicamente por anillos bronquiales, está presente principalmente en Caprimulgiformes (chotacabras), Cuculiformes (cucús) y en algunos Strigidae (lechuzas) [Casey and Gaunt 1985, Elemans et al. 2003, Suthers 2001].

Las siringes traqueales poseen una sola fuente de sonido, como ocurre con la laringe en humanos (*i.e.*, un solo par de cuerdas vocales). Sin embargo, las siringes bronquiales y la mayoría de las traqueobronquiales poseen *dos* fuentes de sonido, situadas una en cada bronquio. Entre las aves que poseen dos, algunas utilizan mayormente un solo lado, como el canario (*Serinus canaria*) [Zollinger and Suthers 2004]; otras utilizan los dos lados, ya sea alternativamente o simultáneamente. Los cuervos (*Molothrus ater*) y los cardenales (*Cardinalis cardinalis*), por ejemplo, usan los dos lados alternativamente [Zollinger and Suthers 2004], mientras que el cuitlacoche rojizo (*Toxostoma rufus*) puede usarlos alternativa o simultáneamente [Goller and Suthers 1996a]. La producción simultánea puede darse con independencia entre ambos lados, como el ya nombrado cuitlacoche [Goller and Suthers 1996a], e inclusive haciéndolos interactuar como en el paro de corona negra [Nowicki and Capranica 1986a].

2.1.2 Mecanismo general de producción de sonido

Debido en parte a la inaccesibilidad de la siringe y a su pequeño tamaño ($\sim 1 - 10$ mm), los primeros intentos de explicar la función de diferentes estructuras en ella fueron puramente conjeturales y se basaban en estudios morfológicos. Más tarde, el desarrollo de delicadas técnicas experimentales permitió establecer con certeza cuál era la fuente de sonido en la siringe.

El mecanismo de fonación en aves es muy similar al de humanos [Fletcher and Tarnopolsky 1999]. Básicamente, la presión del saco aéreo establece un flujo de aire a través de la siringe, que funciona como una válvula vibrante. En ciertas condiciones, la válvula se pone a oscilar en forma espontánea bajo la corriente de aire proveniente del saco aéreo; el movimiento de la válvula modula a su vez la corriente de aire, que es inyectada como una onda de presión acústica en el tracto vocal. La onda acústica es modificada por las resonancias del tracto vocal, en muchos casos como en un sistema fuente-filtro.

En Oscinos, la válvula vocal se forma por la constricción del lumen siríngeo cuando los labia mediales y laterales (ML y LL) obstruyen parcialmente la vía aérea por la acción de los músculos aductores. Este gesto muscular, que lleva la siringe a la *posición fonatoria* [Suthers et al. 1999], es seguido por un aumento en la presión subsiríngea debida a la acción de los músculos expiratorios. El flujo de aire generado produce bajo ciertas condiciones la oscilación de los labia.

La hipótesis de la válvula vibrante no fue el único mecanismo propuesto

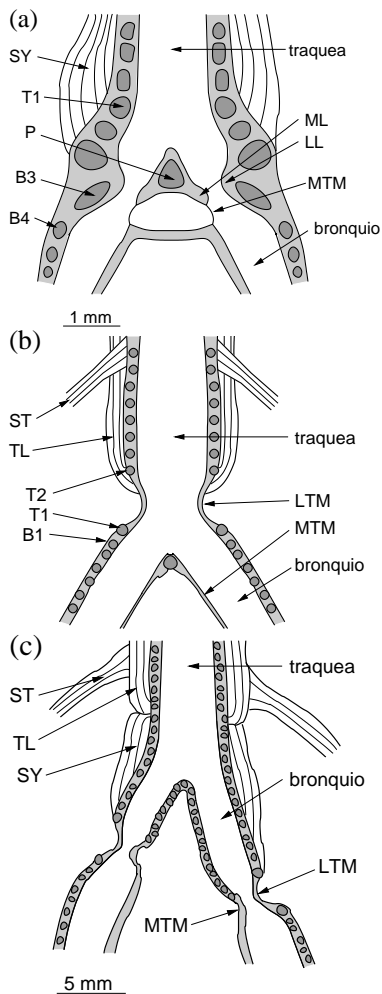


Figura 2.1: Sección ventral de diversas siringes. (a) Siringe traqueobronquial con dos fuentes de sonido, particularmente común en Oscinos (cuitlacoche rojizo, *Toxostoma rufus*). Las dos válvulas vocales se forman por aducción de los labios laterales y mediales de cada lado. LL, ML: labios laterales y mediales; MTM: membranas timpaniformes mediales; T1: primer anillo traqueal; B3, B4: tercer y cuarto anillos bronquiales; P: pessulus, SY: músculo siringeo. (modificado de [Larsen and Goller 2002]). (b) Siringe traqueal (paloma de collar, *Streptopelia risoria*). Una constricción del lumen por las membranas timpaniformes laterales forma la única válvula vocal. LTM: membranas timpaniformes laterales; MTM: membranas timpaniformes mediales; T1, T2: primer y segundo anillos traqueales; B1: primer anillo bronquial; TL: músculo tracheolateralis; ST: músculo sternotrachealis. (modificado de [Gaunt et al. 1982]) (c) Siringe bronquial (guácharo, *Steatornis caripensis*). Notar la diferencia de longitud entre los bronquios. Esta característica deja su huella en el canto de esta ave, cuyo espectro muestra dos formantes diferentes. LTM, MTM: membranas timpaniformes laterales y mediales; TL: músculo tracheolateralis; ST: músculo sternotrachealis; SY: músculo siringeo. (modificado de [Suthers 2001])

para explicar la fonación en aves. Otro mecanismo, el silbido, fue propuesto para dar cuenta de sonidos tonales (esto es, sonidos “puros” con muy pocos armónicos) [Gaunt et al. 1982]. Diversos experimentos, sin embargo, fallaron en dar sustento a esta hipótesis. Las elucubraciones sobre los posibles mecanismos de producción sonora tuvieron su final recién en la última década del siglo XX, con los trabajos de Goller y Larsen [Goller and Larsen 1997a, Goller and Larsen 1997b, Larsen and Goller 1999], quienes mostraron que la fonación se produce en forma análoga a la de humanos, tanto en Oscinos como en loros y en palomas. En todos los casos estudiados, hay evidencia directa de que el mecanismo de fonación es la vibración de estructuras flexibles de la siringe (LTM en palomas y loros, y labia en Oscinos), inclusive cuando se trata de sonidos tonales [Larsen and Goller 1999].

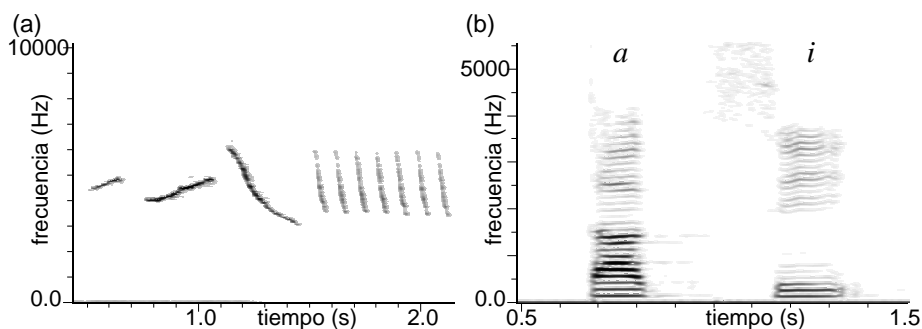


Figura 2.2: Diferencias entre canto de aves y habla humana normal. (a) Canto del chingolo (*Zonotrichia capensis*). Nótese la marcada evolución temporal de la frecuencia fundamental, especialmente en la tercera sílaba que abarca 3000 Hz. La única componente de Fourier presente en esta vocalización es la frecuencia fundamental. (b) Sonograma de la palabra “taxi”. En contraste con el canto del chingolo mostrado a la izquierda, las vocales humanas tienen un gran contenido espectral, mientras que la frecuencia fundamental permanece prácticamente constante.

Aves vs. humanos

El mecanismo básico de fonación en todas las especies de aves estudiadas es similar al de humanos. Hay, sin embargo, notorias diferencias entre el canto de las aves y el habla humana normal: en términos generales, las aves definen las propiedades esenciales de sus vocalizaciones a nivel de la siringe, en tanto que las vocalizaciones normales humanas deben su variedad más que nada a reconfiguraciones del tracto vocal (posición de la lengua, de los labios, etc.).

La Figura 2.2 ilustra este punto. La diversidad de vocalizaciones de aves está dada principalmente por la evolución de la frecuencia fundamental, como resultado del control de los músculos en la fuente sonora. Es cierto que la forma del tracto vocal, el grado de apertura del pico, etc., tienen cierta influencia en la calidad del sonido emitido [Nowicki 1987, Hoese et al. 2000], pero la característica básica es la evolución de la frecuencia fundamental. Por el contrario, en humanos la frecuencia fundamental se mantiene prácticamente constante; la variedad radica en el *timbre*, esto es la cantidad de armónicos presentes en el espectro y sus pesos relativos. Esta propiedad del sonido nos permite reconocer diferentes personas y, más importante aún, nos permite identificar diferentes vocales y por consiguiente interpretar un mensaje. En habla humana normal, el timbre está determinado básicamente por la configuración del tracto vocal, o sea la posición de las diferentes porciones y partes del tracto como la lengua y los labios.

2.2 Oscinos

Las aves que pertenecen al suborden Oscinos tienen la capacidad del aprendizaje vocal. Esta destacable similitud con los humanos hace de los Oscinos un modelo invaluable para estudiar la adquisición del habla y, en general, de un comportamiento complejo. La capacidad de aprender el canto está aparentemente relacionada con un control preciso de los componentes siríngicos. En esta

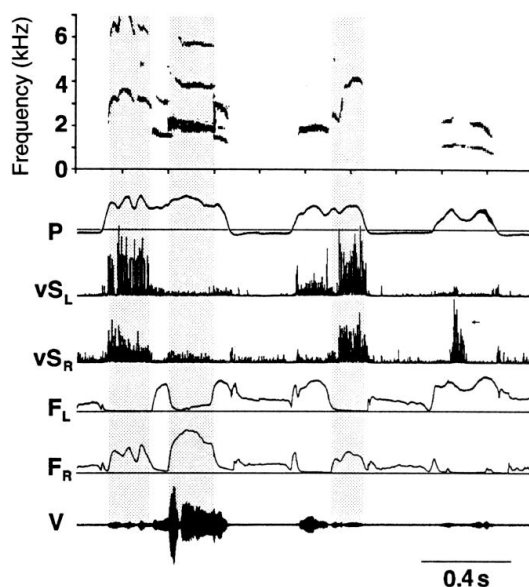


Figura 2.3: Mediciones de actividad muscular (vS), presión del saco aéreo (P), flujo de aire a través de los bronquios (F) y vocalización (V + sonograma), todo en simultáneo y en ambos lados de la siringe (L = izquierda, R = derecha), en un cuitlacoche rojizo despierto y cantando espontáneamente. (extraído de [Goller and Suthers 1995])

sección describiremos la delicada complejidad de su aparato vocal.

2.2.1 El rol de los músculos

Con anterioridad a que el origen preciso del sonido en la siringe fuera determinado, Goller y Suthers ya habían estudiado —con un detalle inédito hasta ese momento— el rol de cada músculo de la siringe en las vocalizaciones de un cuitlacoche rojizo *in vivo* [Goller and Suthers 1995, Goller and Suthers 1996a, Goller and Suthers 1996b]. Esta tarea es aún más difícil que la determinación del origen del sonido. El estudio del rol muscular requiere la medición directa de actividad electromiográfica de cada uno de los músculos que gobiernan la siringe mientras el ave está vocalizando, y consiste en insertar un electrodo en el músculo en cuestión. El voltaje registrado es una medida de la actividad muscular. Es necesario, además, medir la presión en el saco aéreo y el flujo de aire a través de los bronquios. Lo primero se logra insertando una cánula conectada a un transductor de presión; lo segundo, con un termistor insertado en cada bronquio. La medición simultánea de presión, flujo de aire, actividad muscular en varios músculos y vocalización (Figura 2.3) permitió desentrañar el papel de cada músculo en la siringe.

En la Figura 2.4 se muestra esquemáticamente la disposición de los músculos en la siringe. Los resultados de Goller y Suthers indican en principio que la fonación se da siempre con flujo de aire a través del bronquio, generado por una presión en el saco aéreo mayor que la atmosférica. En lo que se refiere a los músculos siríngeos, el papel de los músculos syringealis dorsalis (dS) y tracheo-bronchialis dorsalis (dTb) es de aducción, o sea cierre activo del lumen siríngeo.

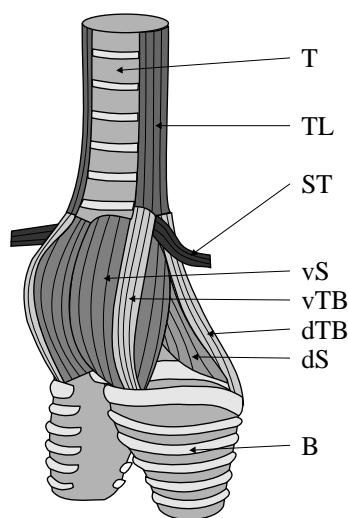


Figura 2.4: Vista ventrolateral esquemática de una siringe de Oscino. Los músculos (o grupos de músculos) tienen roles más o menos bien definidos, como por ejemplo el control de la frecuencia de la vocalización o el cierre y la apertura activos del lumen siríngeo. vS: músculo siringealis ventralis; dS: m. siringealis dorsalis; vTB: m. tracheobronchialis ventralis; dTB: m. tracheobronchialis dorsalis; TL: m. tracheolateralis; ST: m. sternotrachealis; B: bronquio, T: tráquea. (modificado de [Larsen and Goller 2002])

Es decir, el ave puede impedir activamente el paso de aire. Por otro lado, la apertura activa del lumen, o abducción, está a cargo de los músculos tracheolateralis (TL) y tracheobronchialis ventralis (vTB). A partir de los experimentos mencionados quedó demostrado que el ave utiliza los músculos aductores para gobernar con precisión el comienzo y el fin del sonido. Inmediatamente después de una vocalización utiliza los músculos abductores para abrir activamente el lumen y dar lugar rápidamente a una inspiración [Goller and Suthers 1996a].

Mientras dura la vocalización, la frecuencia del sonido emitido parece estar gobernada por el músculo más grande de la siringe, el siringealis ventralis (vS). La actividad de este músculo está altamente correlacionada con la frecuencia fundamental de la vocalización: su actividad se incrementa exponencialmente con la frecuencia del sonido [Goller and Suthers 1995, Goller and Suthers 1996b] (Figura 2.5). El rol de este músculo es común en varias especies de Oscinos, como cuitlacoche rojizo, cardenal, cuervo y canario [Suthers et al. 1999]. Finalmente, el músculo sternotrachealis (ST) no tiene un rol aparente tan diferenciado; se cree que cumple el papel de estabilizar la estructura siríngea para permitir la rápida acción de los músculos aductores y abductores [Goller and Suthers 1996a].

Aunque es probable que en algunas aves el rol de los músculos no esté tan diferenciado, como por ejemplo el pinzón cebra (*Taenopygia guttata*, donde el vS tendría algún papel en la abducción [Vicario 1991]) o el cardenal (donde aparentemente los músculos dorsales también influyen en la frecuencia de la vocalización [Mindlin et al. 2003]), los roles descritos más arriba parecen ser comunes en Oscinos [Suthers 2001].

Las riquísimas mediciones de Goller y Suthers, sin embargo, no fueron aprovechadas para desarrollar modelos con sustento experimental hasta pasados algunos años.

2.2.2 Lateralización

La identidad bipartita de la fuente sonora en Oscinos nos obliga a considerar la simetría del control motor. ¿Qué grado de independenciamos tienen los lados de la

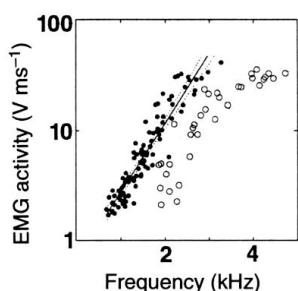


Figura 2.5: Relación entre actividad muscular en el músculo vS y frecuencia fundamental del sonido emitido (círculos negros) en un cuitlacoche rojizo. (extraído de [Goller and Suthers 1995])

siringe? Hemos dicho ya que muchas especies de aves utilizan ambos lados de la siringe en mayor o menor medida. Los canarios de raza Waterslager, en un extremo de la dominancia lateral, cantan alrededor del 90% de las sílabas con el lado izquierdo, mientras que en el otro extremo el cuitlacoche rojizo canta con aportes parejos de ambos lados [Suthers et al. 1999].

Para empezar, no hay evidencia de lateralización en la actividad respiratoria durante el canto. Mediciones de actividad electromiográfica de los músculos expiratorios abdominales izquierdo y derecho en el cuitlacoche rojizo mostraron que tanto el patrón temporal como la amplitud media de la actividad muscular son muy similares, sin importar qué lado de la siringe está fonando [Goller and Suthers 1999]. De todas formas, dado que los sacos aéreos de cada lado están interconectados a través del saco aéreo interclavicular y otros sacos, es dudoso que la actividad unilateral de los músculos respiratorios (si ocurriera) pueda tener algún efecto sobre un solo lado de la siringe [Suthers 1997, Suthers et al. 1999].

En cambio, los músculos siríngicos dorsales (dS y dTB), que gobiernan el cierre activo del lumen siríngeo para permitir o impedir la fonación del lado correspondiente, son herramientas primordiales para ejecutar los diferentes patrones de lateralización descritos recién. Los programas motores enviados a estos músculos son los que determinan qué lado de la siringe debe fonar [Suthers et al. 1999].

El músculo que gobierna la frecuencia del sonido (vS) también recibe diferentes programas motores de cada lado, dando lugar a vocalizaciones que tienen simultáneamente dos frecuencias no relacionadas armónica ni temporalmente [Suthers et al. 1999], fenómeno conocido tradicionalmente como “dos voces”. Sorprendentemente, este músculo está activo sin embargo en ambos lados de la siringe, aunque uno de los lados esté silenciado por la acción de los aductores. Esta observación sugiere que el patrón motor fonatorio del lado silenciado es ejecutado a pesar de estar cerrado por los aductores. Siendo el vS el mayor músculo de la siringe, la actividad en el lado silenciado podría ser ventajosa, al estabilizar el marco siríngeo contra el cual los músculos contralaterales del lado que está fonando deben actuar [Suthers et al. 1999].

Este control unilateral periférico de la expresión de patrones motores bilaterales podría explicar la ausencia de asimetría hemisférica entre los núcleos neuronales del control motor, tanto en morfología como en actividad, así como la paradoja de una producción de canto altamente asimétrica en algunas especies [Goller and Suthers 1995, Suthers 1997].

Otra posible ventaja de la actividad continua del músculo vS en ambos lados

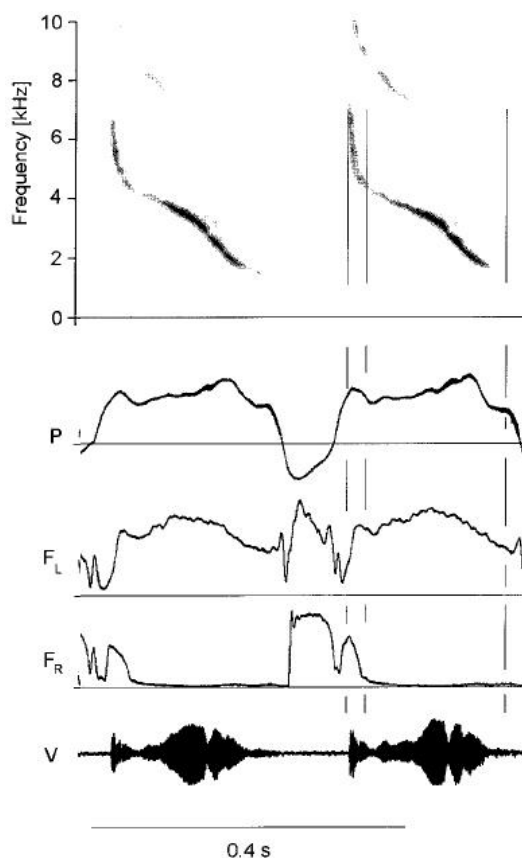


Figura 2.6: Sílabas de cardinal realizadas con ambos lados de la siringe actuando secuencialmente. La parte con frecuencias por encima de 4000 Hz es realizada por el lado derecho; la parte con frecuencias por debajo de 4000 Hz, por el lado izquierdo. La continuidad en la transición es asombrosa. P: presión en el saco aéreo; F_L, F_R: flujo en los lados izquierdo y derecho; V: envolvente del sonido. Extraído de [Suthers 1997].

de la siringe (a pesar de que uno de ellos esté silenciado) sería que el sonido pueda ser iniciado casi instantáneamente a la frecuencia correcta cuando el flujo de aire cambie abruptamente de lado [Suthers 1997]. El canto de los cardenales, por ejemplo, se caracteriza por sílabas que cubren en forma continua el rango que va de los 2000 Hz hasta los 7000 Hz, monótonamente hacia arriba o hacia abajo (ver Figura 2.6). La parte de la sílaba con frecuencias por encima de 4000 Hz es ejecutada siempre por el lado derecho, mientras que la parte con frecuencias por debajo de 4000 Hz es ejecutada por el lado izquierdo, con una transición suave y coordinada entre ambos lados en el medio de la sílaba [Suthers 1997]. Este *tour de force* de la coordinación motora izquierda-derecha se contrapone con el hecho de que los centros neuronales de control del canto tienen sólo proyecciones limitadas al lado contralateral del cerebro, y no se conoce en profundidad qué conexiones median la coordinación motora entre los lados de la siringe [Suthers 1997, Wild 1997]. Una pregunta es relevante, entonces: ¿es la física del aparato fonador una aliada en esta coordinación? Las condiciones de oscilación de un lado de la siringe a una cierta frecuencia podrían verse favorecidas por el hecho

de que el otro lado se encuentra oscilando a una frecuencia similar.

2.2.3 Aves que aprenden y musculatura intrínseca

Los músculos que gobiernan la siringe pueden clasificarse en *intrínsecos* y *extrínsecos*. Los primeros se insertan enteramente dentro de la siringe, y los segundos nacen o terminan fuera de la siringe. Los intrínsecos probablemente controlan la posición relativa de las partes de la siringe, mientras que los extrínsecos mueven la siringe como un todo [Gaunt 1983, Suthers 2001, Larsen and Goller 2002].

Es interesante notar que los músculos extrínsecos son comunes a todas las aves, pero no todas tienen además músculos intrínsecos. Es más, la capacidad de aprendizaje vocal parece restringida a especies con musculatura intrínseca, independientemente del número de músculos intrínsecos. Loros y colibríes tienen 2 pares, mientras que los Oscinos tienen por lo menos 4 pares, como la calandria (*Mimus polyglottos*). Ha sido propuesto que el control independiente de los componentes siríngicos sería una condición necesaria (aunque no suficiente) para el aprendizaje vocal, dado que libera las restricciones mecánicas de una siringe simple [Gaunt 1983].

2.3 No-Oscinos

Mientras que los Oscinos se caracterizan por su siringe traqueobronquial y bilateral, la siringe de los no-Oscinos es mucho más variada. En algunos casos es una siringe traqueobronquial, pero con una única válvula vocal; en otros casos es enteramente traqueal, y hay casos en los que es bronquial. En todos los casos, como ya dijimos, el mecanismo de fonación es el mismo: la siringe posee tejidos móviles que obstruyen parcialmente el paso del aire, y que en ciertas condiciones se ponen a oscilar en forma autosostenida modulando el flujo de aire hacia el tracto vocal. Entre Oscinos y no-Oscinos hay, sin embargo, profundas diferencias en lo que se refiere al control muscular de la siringe y por ende de las propiedades del sonido emitido. En lo que sigue describiremos algunas siringes de no-Oscinos y discutiremos brevemente estas diferencias.

Entre los que no aprenden: palomas

Como muestra del aparato fonador de un no-Oscino que no tiene la capacidad de aprendizaje vocal, en la Figura 2.1b se ve una siringe de paloma. Es una siringe traqueal, y por lo tanto tiene una sola válvula vocal, formada por la aducción de las LTM. La configuración de la siringe está determinada por la acción de dos músculos, ambos extrínsecos: el tracheolateralis (TL), que se inserta directamente sobre las LTM, y el sternotrachealis (ST), que se inserta en la tráquea. La contracción del TL causa la abducción de las LTM; eventualmente, puede apartarlas totalmente del lumen siríngico y terminar la fonación [Goller and Larsen 1997a]. La contracción del ST mueve la tráquea en dirección caudal (hacia la cola) y causa el plegamiento de las LTM hacia el lumen siríngico. Sin embargo, la máxima contracción del ST aparentemente no alcanza para cerrar el conducto [Goller and Larsen 1997a].

La posición de las LTM cambia al vocalizar. La presión del saco aéreo interclavicular provoca el avance de las LTM sobre el lumen siríngico, pues la

superficie externa de las LTM está en contacto directo con el espacio de aire del saco interclavicular. Una vez que las membranas están completamente aducidas comienza la fonación. La contracción leve del TL provoca cambios en el sonido, ya sea una disminución en la frecuencia fundamental o cambios en la estructura de armónicos [Goller and Larsen 1997a, Elemans et al. 2004]. Una observación interesante es que la tensión de las LTM puede estar gobernada, además, por la presión del saco aéreo interclavicular: la frecuencia del sonido y la presión del saco están correlacionadas positivamente [Beckers et al. 2003b]. La evidencia indica que estas aves utilizan mecanismos diferentes que los Oscinos para controlar la oscilación de la válvula vocal [Suthers 2001].

Entre los que aprenden: loros

Como dijimos, la mayoría de las aves que aprenden pertenecen al suborden Oscinos. El resto de las aves que aprenden son loros y colibríes. A continuación describiremos la siringe de los loros.

Aunque es traqueobronquial, la siringe en loros tiene una única válvula vocal formada por las membranas timpaniformes laterales (LTM). Las LTM se insertan entre el primer anillo traqueal y el primer anillo bronquial, y se encuentran plegadas hacia el lumen siríngeo como resultado de la proximidad entre dichos anillos cartilagosos [Larsen and Goller 2002]. Como fue descrito en el caso de palomas, la superficie lateral (“externa”) de las LTM está expuesta a las fluctuaciones de la presión del saco aéreo interclavicular, y pueden ocluir el lumen siríngeo o retraerse según sea la presión en el saco [Suthers and Zollinger 2004].

Tres músculos son los que determinan la configuración de la siringe: el extrínseco sternotrachealis (ST), y los intrínsecos syringealis superficialis (SS) y syringealis profundus (SP).¹ A pesar de que los músculos intrínsecos le dan un control fino de la configuración de la siringe, la regulación de la tensión de las LTM en loros parece ser más indirecta que en Oscinos [Larsen and Goller 2002], pues la frecuencia del sonido no está fuertemente correlacionada con la actividad de ninguno de los músculos de la siringe [Gaunt and Gaunt 1985]. Las similitudes anatómicas entre las siringes de palomas y loros merecen continuar la investigación sobre los mecanismos utilizados por dichas especies para controlar del sonido [Suthers and Zollinger 2004].

¹El delgado músculo extrínseco tracheolateralis (TL) parece no tener un papel preponderante en la fonación [Larsen and Goller 2002].

Capítulo 3

Fonación

El mecanismo y las estructuras siríngicas responsables de la fonación fueron temas controversiales durante mucho tiempo. En este capítulo haremos primero un repaso sobre la historia de los modelos de producción de sonido en la siringe, y luego describiremos los experimentos cruciales que permitieron armar el rompecabezas de la fonación en aves. Luego presentaremos los modelos básicos de fonación desarrollados en esta tesis. Finalizaremos este capítulo con una discusión sobre el papel del tracto vocal en el canto de las aves.

3.1 Un poco de historia

Una reseña histórica sobre la investigación de los mecanismos sonoros en la siringe tiene que empezar, casi forzosamente, en una variedad de modelos hipotéticos propuestos para su funcionamiento, antes que en la poca evidencia disponible en aquellos tiempos. El pequeño tamaño de la siringe y su ubicación (profunda en el tórax, dentro de un saco aéreo y cerca del corazón) siempre fueron inconvenientes para observar la función de diferentes partes de la siringe, o inclusive para arriesgar modelos matemáticos. Por esta razón, al comienzo los modelos para la fonación en aves fueron puramente conceptuales. Las técnicas experimentales que permitieron tener evidencia directa del mecanismo de fonación fueron desarrolladas en forma relativamente reciente (en la última década del siglo XX).

3.1.1 Modelos vibrantes

En el año 1951, Miskimen publicó una observación experimental que habría de ser la base del primer mecanismo propuesto para la fonación en la siringe. Miskimen observó que, cuando hacía pasar aire a través de la siringe de un gorrión (*Passer domesticus*), ya sea inyectando aire en el saco aéreo o succionando a través de la tráquea, se producía un sonido “como de gorjeo” mientras las MTM vibraban [Miskimen 1951]. Concluyó que el sonido era producido por la vibración de las delgadas MTM, mecanismo que fue modelado conceptualmente como una membrana vibrante acoplada al flujo de aire [Greenewalt 1968]. Este mecanismo llegó a ser conocido como el “modelo clásico” de fonación en aves.

Otro modelo, también conceptual, propuso para la fonación de los Oscinos

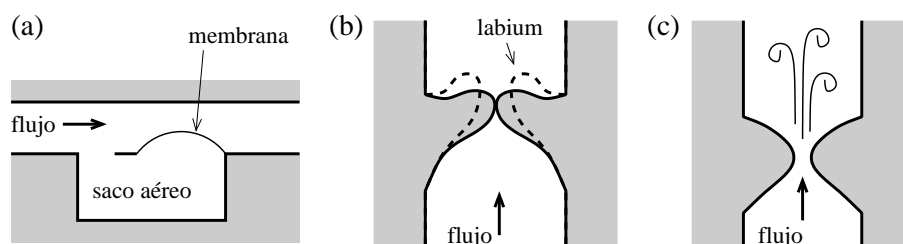


Figura 3.1: Representación esquemática de los principales modelos conceptuales para la fonación en aves. (a) Modelo de membranas o “modelo clásico”. El movimiento de una membrana delgada acoplada al flujo de aire produce cambios acústicos en la densidad del flujo, generando sonido. (b) Modelo de válvula vibrante gobernada por presión. El flujo de aire provoca la apertura periódica de la válvula; ésta a su vez modula el flujo de aire, que es inyectado en el tracto vocal en forma pulsátil. (c) Modelo aerodinámico (“silbido”). Una constricción en la siringe forma una ranura angosta donde se produce desprendimiento de vórtices y turbulencia o *ruido eólico*. La interacción con las resonancias del tracto acentúa una de las frecuencias presentes, como en una flauta o un silbato.

un mecanismo tipo válvula gobernada por presión análogo al de las cuerdas vocales humanas. Este modelo [Klatt and Stefanski 1974], basado en observaciones anatómicas *in vitro* —y en una sorprendente intuición—, sostiene que la fuente de sonido son los más gruesos labia laterales (LL). Además, propone para los LL un movimiento “tipo onda” completamente análogo al que les permite a las cuerdas vocales humanas oscilar en forma autosostenida [Titze 1988]. Una representación esquemática de los modelos puede verse en la Figura 3.1.

3.1.2 La hipótesis del silbido

La mayor dificultad que tuvieron los modelos vibrantes para lograr su aceptación fue su aparente imposibilidad de producir sonidos armónicos (con componentes armónicas) y sonidos puros (carentes de sobretonos, llamados también sonidos *tonales*), que aparecen en muchas vocalizaciones de aves.

Las críticas a los modelos de membrana se basaban en que una membrana aproximadamente redonda y fija por su borde como las MTM no produciría sonidos armónicos sino *anarmónicos*: las componentes del sonido estarían dadas por los ceros de una función de Bessel, y por lo tanto no serían múltiplos enteros de la fundamental. Por otro lado, en un mecanismo tipo válvula de presión, donde las estructuras vibrantes se chocan una contra otra como las cuerdas vocales humanas, el sonido producido podría ser armónico pero tendría muchas componentes como resultado del movimiento no sinusoidal de la fuente. El conocimiento previo del mecanismo de fonación en humanos jugó aquí una mala pasada: la supresión casi total de armónicos por filtrado en el tracto vocal era considerada poco probable.

Una de las alternativas que surgió para resolver estos dilemas fue la *hipótesis del silbido*, reflejada en los llamados “modelos aerodinámicos” [Gaunt et al. 1982, Casey and Gaunt 1985]. Estos modelos descartan cualquier vibración de alguna parte de la siringe y proponen que ésta funciona como un silbato. En este mecanismo, el paso del aire a través de una “ranura” formada por un

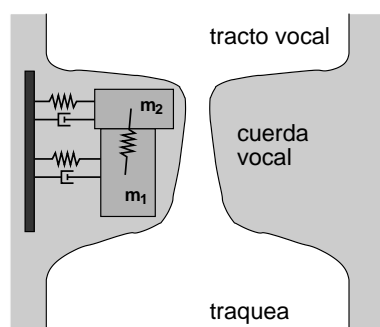


Figura 3.2: Modelo de dos masas para la oscilación de las cuerdas vocales humanas. Este modelo reproduce en forma espontánea la diferencia de fase observada experimentalmente entre las partes superior e inferior de la cuerda vocal. Esta diferencia de fase es la que permite una transferencia neta de energía del flujo de aire a las cuerdas vocales en un ciclo.

angostamiento de la siringe produce turbulencia; la posterior interacción con las resonancias del tracto vocal acentúa mayormente una de las componentes espectrales presentes, como en una flauta o un silbato.

Aún así, hasta la última década del siglo XX los modelos cuantitativos de fonación en aves eran escasos: de membranas [Brackenbury 1979]; de membranas, modificado [Casey and Gaunt 1985, Fletcher 1988]; válvula de presión [Fee et al. 1998, Fry 1998].

3.1.3 Modelos de fonación en humanos

La falta de datos experimentales en aves hasta épocas recientes explica la escasez de modelos matemáticos que describan el mecanismo sonoro de la siringe. Este hecho contrasta con el caso de fonación en humanos, donde la abundancia de modelos cuantitativos puede verse simbólicamente reflejada, por ejemplo, en el sorprendente número de masas utilizadas para modelar el tejido cordal. La lista empieza con modelos de una sola masa con acople a un tracto vocal [Flanagan and Landgraf 1968], y sigue con modelos de dos masas [Ishizaka and Flanagan 1972], de tres masas [Story and Titze 1995], de 16 masas [Titze 1973, Titze 1974], de $3 \times 3 \times 3$ masas [Titze and Story 1997], y hasta modelos de medio continuo [Berry and Titze 1996]. Otros trabajos abandonaron la idea de reproducir los detalles biomecánicos de las cuerdas vocales y se inclinaron por desarrollar modelos basados en parámetros glotométricos (como el área de contacto entre las cuerdas y el área de la glotis, extraídos de glotografías) [Titze 1984].

Uno de los primeros y más exitosos modelos de vibración de las cuerdas vocales humanas es el conocido “modelo de dos masas” de Ishizaka y Flanagan [Ishizaka and Flanagan 1972, Steinecke and Herzel 1995]. En este modelo se asume que las cuerdas vocales son inhomogéneas en estructura, con dos partes móviles diferentes acopladas entre sí en forma elástica (Figura 3.2). La elección de esta geometría tiene una razón: técnicas endoscópicas revelan que en el movimiento oscilatorio de la cuerda vocal se observa una diferencia de fase entre sus bordes superior e inferior [Titze 1988, Švec et al. 1996]. Esta diferencia de fase es la que permite una transferencia neta de energía del flujo de aire a las cuerdas vocales. Si la parte superior de la cuerda está retrasada con respecto a la parte inferior, la presión entre las cuerdas es mayor (más parecida a la presión traqueal) en la fase de apertura que en la fase de clausura [Titze 1988, Laje et al. 2001]. El modelo de dos masas reproduce espontáneamente esta diferencia de fase.

A mitad de camino entre los modelos de una y de dos masas, un ingenioso modelo propuesto por Titze [Titze 1988] *impone* la diferencia de fase entre los bordes de la cuerda vocal a través de una geometría asumida para el perfil de la cuerda, lo que permite escribir una sola ecuación diferencial de segundo orden. Este modelo, descrito en la Sección 3.3.3, constituye la base de los modelos desarrollados en esta tesis.

3.2 Experimentos definitivos

Durante mucho tiempo, la inaccesibilidad y el pequeño tamaño de la siringe fueron escollos para determinar experimentalmente la validez de los modelos propuestos para la fonación. El camino hacia la validación de los modelos fue iniciado con métodos experimentales indirectos, como la grabación de vocalizaciones en una atmósfera de heliox. Las pruebas experimentales directas se hicieron esperar hasta la última década del siglo XX.

3.2.1 Experimentos con heliox

Los primeros experimentos para validar los modelos de fonación consistieron en grabar vocalizaciones de aves en un atmósfera de *heliox*. El heliox es una mezcla en la cual el nitrógeno, que compone el 80% del aire, es reemplazado por el menos denso helio. El resultado es que la velocidad del sonido es 74% mayor que en aire [Gray 1963, Nowicki 1987]. Los experimentos con heliox, que son relativamente simples, permitieron someter a validación el mecanismo del silbido.

La frecuencia fundamental de un silbato está determinada básicamente por la frecuencia natural del resonador acoplado. Muchos instrumentos musicales de viento, como la flauta, funcionan de la misma manera. Una corriente de aire se hace pasar a través de un angostamiento (una ranura, un agujero, o inclusive un borde filoso), lo que rompe la suavidad del flujo y produce turbulencia. La turbulencia y el desprendimiento de vórtices excitan la columna de aire de la cavidad acoplada con muchas frecuencias. La interacción con las resonancias de la cavidad hace que una única frecuencia sea amplificada: la que coincide con la frecuencia de resonancia.

Por otro lado, la frecuencia fundamental de resonancia de un tubo es directamente proporcional a la velocidad del sonido. Por lo tanto, si los sonidos tonales en aves fueran producidos como en un silbato, la frecuencia de un sonido emitido en una atmósfera menos densa debería aumentar en la misma proporción en que aumenta la velocidad del sonido. En cambio, si el mecanismo de fonación fuera similar al de humanos, la frecuencia fundamental del sonido emitido no debería cambiar. En el caso de fonación en humanos, el helio no tiene influencia en la frecuencia fundamental de oscilación de las cuerdas vocales [Beil 1962], pues la vibración de las cuerdas vocales es en gran medida independiente de las resonancias del tracto vocal (en habla normal [Laje et al. 2001]).

En todas las especies estudiadas, la frecuencia fundamental de las vocalizaciones no se modificó sustancialmente al cambiar a una atmósfera de heliox. Los experimentos permitieron descartar la hipótesis del silbido en Oscinos [Nowicki 1987], en loros [Brittan-Powell et al. 1997] y en palomas [Ballintijn and ten Cate 1998].

3.2.2 Endoscopía y lesiones

Los experimentos de heliox refutaron la hipótesis del silbido y aportaron evidencia indirecta a favor de un mecanismo vibratorio. Sin embargo, experimentos de este tipo no permiten discernir entre las diferentes fuentes propuestas para la vibración, *e.g.* los labia laterales y mediales (LL y ML) vs. las membranas (MTM) en Oscinos, y las membranas LTM vs. MTM en palomas.

Los primeros experimentos para determinar directamente el origen del sonido en la siringe consistieron en una combinación de observaciones endoscópicas *in vivo* y estudios de lesiones a nivel de la siringe. Goller y Larsen mostraron que la producción de sonido estaba siempre acompañada por vibraciones de ciertas estructuras en la siringe: los labia LL y ML en Oscinos [Goller and Larsen 1997b], y las LTM en palomas [Goller and Larsen 1997a]. Además, la incapacitación de las MTM en todas las especies estudiadas no evitó la fonación, que se producía en forma prácticamente normal [Goller and Larsen 1997a, Goller and Larsen 1997b].

Los resultados de estos experimentos sugirieron fuertemente que la fonación se produce por un mecanismo vibratorio, y que las MTM no son la fuente principal de sonido. Sin embargo, la baja resolución temporal del sistema endoscópico no permitió un análisis cuantitativo de las vibraciones observadas [Larsen and Goller 1999]. En un experimento mejorado, Larsen y Goller utilizaron un detector óptico para medir la vibración en la siringe en un Oscino, una paloma y un loro, y determinaron que la frecuencia de oscilación de los tejidos coincidía con la frecuencia del sonido emitido en todos los casos [Larsen and Goller 1999].

En resumen, ahora sabemos que en todas las especies estudiadas el mecanismo de fonación es similar al de humanos. Un par de masas de tejido obstruyen parcialmente el lumen siringeo y actúan como una válvula vibrante, que bajo ciertas condiciones se abre y se cierra en forma espontánea ante el flujo de aire proveniente de los sacos aéreos. Este movimiento modula a su vez al flujo de aire que es inyectado como una onda de presión en el tracto vocal.

3.3 Oscilaciones en la siringe

Los experimentos de Goller y Larsen descritos en la sección anterior dejaron bien establecido que el mecanismo de fonación en aves es similar al de humanos. Por esta razón, los modelos desarrollados para cuerdas vocales humanas pueden ser adaptados al canto de las aves.

Los detallados experimentos de Goller y Larsen de la sección anterior y los de Goller y Suthers descritos en la Sección 2.2.1 permitieron el desarrollo de modelos de fonación basados en un sólido conocimiento de la mecánica de la siringe. Dichos modelos, que son los desarrollados en esta tesis, son presentados en las siguientes dos secciones.

3.3.1 Por qué modelar la siringe

Una gran cantidad de trabajo está dedicado actualmente a entender cómo se representan en el cerebro del ave la percepción y la producción del canto [McCasland 1987, Brainard and Doupe 2002]. El desarrollo de sofisticadas técnicas de electrofisiología —por ejemplo, la medición simultánea de la actividad de varias neuronas individuales *in vivo* [Hahnloser et al. 2002]— y de expresión

genética [Jarvis et al. 2002] permitió evaluar una cantidad de hipótesis sobre el rol de diferentes partes del cerebro. Sin embargo, estas aproximaciones “desde arriba” se enfrentan siempre con la dificultad de interpretar la abrumadora cantidad de datos que surgen de mediciones de la actividad neuronal central (aún en el caso de las aves canoras donde los centros neuronales del canto están altamente especializados [Farries 2001] y son, por lo tanto, más accesibles al análisis).

En contraposición a la tendencia “cerebral” de los estudios de canto de pájaros, el primer modelo cuantitativo [Fee et al. 1998] publicado después de las detalladas observaciones de Goller y Suthers atacó el problema “desde abajo”. Este trabajo demostró que las características acústicas de una vocalización no dependen únicamente del control neuronal, sino que puede haber un grado adicional de complejidad espectral y temporal dado puramente por las propiedades mecánicas no lineales de la siringe.

Queda claro que, para establecer relaciones entre las propiedades del sustrato neuronal y el comportamiento observado, es necesario conocer el mecanismo de interfase entre sustrato y comportamiento: es indispensable una comprensión física del fenómeno de producción de sonido en el órgano vocal del ave. A fin de relacionar las actividades neuronales con el sonido emitido, un conocimiento de la mecánica del aparato productor de sonido es esencial. Esta aproximación “desde abajo” es novedosa y puede aportar mucho al entendimiento del control motor, ayudando a establecer un puente entre actividad motora central y periférica [Suthers and Margoliash 2002, Suthers and Zollinger 2004].

3.3.2 Oscilaciones autosostenidas

La condición para que se produzca la oscilación autosostenida en la válvula vocal es que exista una transferencia neta de energía del flujo de aire a las partes móviles de la válvula, de manera de sobrepasar las pérdidas por disipación en un ciclo de la oscilación.

En términos generales, si un sistema está sujeto a una fuerza habrá una transferencia de energía al sistema si la fuerza aplicada tiene una componente que apunta en el mismo sentido que la velocidad del sistema. En el caso de un sistema oscilante, la fuerza aplicada (también oscilante) debe tener una componente en fase con la velocidad del sistema para que exista transferencia neta en un ciclo. De otro modo la energía suministrada en una parte del ciclo es quitada en la parte del ciclo donde la fuerza es opuesta a la velocidad.

En el caso de la válvula vocal, la fuerza responsable de entregar energía a las partes móviles proviene de la presión de aire que se crea en el interior de la válvula. La transferencia de energía del flujo de aire al tejido móvil es posible cuando dicha fuerza tiene una componente en fase con la velocidad de los tejidos. La clave está en entender cómo se puede lograr una fuerza aerodinámica con tal característica.

Queda claro que debe existir una asimetría entre ambas mitades del ciclo de oscilación, pues la consideración ingenua del teorema de Bernouilli $p_s - p_g = \frac{1}{2}\rho(v_g^2 - v_s^2)$ entre la región sublabial s y la válvula g resulta en que todo lo que se gana en la fase de apertura se pierde en la fase de cierre [Titze 1988].

Acople inercial

La necesaria asimetría puede darse por diferentes causas. Consideremos primero el caso de acople inercial al tracto vocal: el aire que sale de la válvula se encuentra con la columna de aire en el tracto, que en la aproximación de bajas frecuencias se mueve como un todo. Esta hipótesis de acople inercial al tracto vocal es la base del modelo de una masa [Flanagan and Landgraf 1968, Titze 1988]. La inercia de la columna de aire provoca que la presión en la válvula sea mayor en la fase de apertura (cuando el flujo aumenta) que en la fase de cierre (cuando el flujo disminuye), lo que significa que en un ciclo el aporte neto de la fuerza está en el sentido de la velocidad del tejido.

Punto móvil de separación del *jet*

Aunque la hipótesis de acople inercial al tracto vocal puede cumplirse en ciertas condiciones, experimentos con laringes y siringes aisladas (extraídas quirúrgicamente) muestran que la oscilación se produce aún sin acople a un tracto vocal. Es necesario entonces encontrar otras causas para la asimetría mencionada arriba. Una de las explicaciones posibles se basa en el hecho de que el flujo a través de la válvula se separa de las paredes y forma un *jet* si la sección del canal aumenta abruptamente, como en el caso de una válvula con perfil altamente divergente. En cambio, el flujo es laminar si la válvula tiene perfil convergente. Está demostrado que el movimiento del punto de separación tiene importancia en las condiciones de oscilación de las cuerdas vocales humanas [Pelorson et al. 1994, Lucero 1998]. El movimiento del punto de separación del *jet* entre las fases de apertura y cierre es lo que da asimetría a la fuerza aerodinámica.

Movimiento flameante

Un mecanismo alternativo para lograr que la fuerza aerodinámica sea asimétrica dentro de un ciclo de oscilación fue propuesto por Titze en un modelo de las cuerdas vocales humanas [Titze 1988]. El modelo asume que el movimiento de las cuerdas vocales tiene dos componentes: la oscilación lateral de su centro de masa, y una onda viajera en la superficie del tejido. Si estas componentes se acoplan de manera de lograr un movimiento “flameante” (es decir, tal que el perfil de la válvula sea convergente cuando las cuerdas se están separando y divergente cuando se están aproximando), como se ve en la Figura 3.3, la presión es mayor en la fase de apertura y menor en la fase de cierre.

La hipótesis del movimiento “flameante” es consistente con observaciones videoscópicas y videostroboscópicas de una onda viajera en la superficie del tejido, tanto en humanos [Titze 1988, Švec et al. 1996] como en aves [Fee et al. 1998, Larsen and Goller 1999]. Esta onda viajera se debe posiblemente a desplazamientos de la mucosa y es típicamente referida como una diferencia de fase entre las porciones inferior y superior de la cuerda vocal o el labium.

3.3.3 Modelo flameante de Titze

Con el fin de estudiar los principios básicos que rigen la oscilación de las cuerdas vocales humanas, Titze propuso un modelo simple [Titze 1988] a mitad de camino entre el modelo de una masa y el de dos masas. Este modelo supone la existencia de una onda viajera hacia arriba en la superficie de la cuerda vocal,

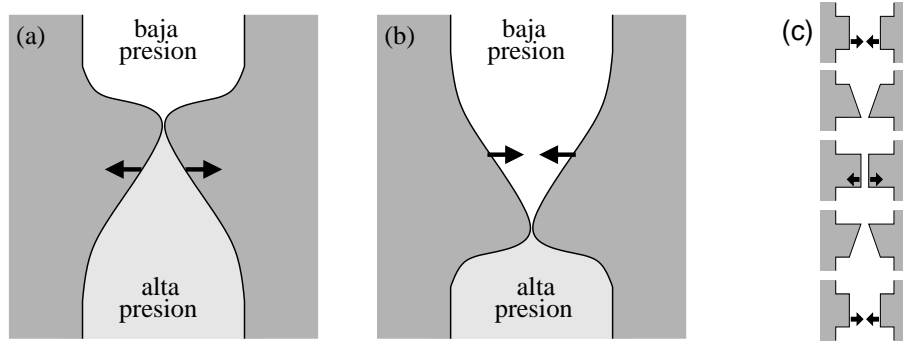


Figura 3.3: Mecanismo de transferencia de energía del flujo de aire al movimiento de la válvula fonatoria. Si la válvula tiene un perfil convergente en la fase de apertura y un perfil divergente en la fase de cierre, la presión que separa la válvula es mayor en la fase de apertura que en la fase de cierre. Esta asimetría permite la transferencia neta de energía en un ciclo de oscilación. (a) Fase de apertura, perfil convergente. (b) Fase de cierre, perfil divergente. (c) De arriba hacia abajo, secuencia temporal esquemática del movimiento de flameo en un ciclo de oscilación.

además de la oscilación lateral del centro de masa de la cuerda. La composición de estos dos movimientos en la forma de un “flameo” (es decir, tal que las cuerdas presenten un perfil convergente cuando se separan y un perfil divergente cuando se aproximan) resulta en una presión intraglotal mayor en la fase de apertura que en la fase de cierre, como fue discutido en la subsección anterior.

La hipótesis de la onda viajera en la superficie de la cuerda vocal está basada en evidencia experimental [Titze 1988, Švec et al. 1996]. Estudios videostroboscópicos de las cuerdas vocales, tanto *in vivo* como en laringes extraídas quirúrgicamente, mostraron un movimiento global tipo flameo. La onda viajera superficial observada se debe probablemente a un desplazamiento de la mucosa que se propaga hacia arriba.

El modelo flameante de Titze, que fue originalmente propuesto para estudiar oscilaciones de pequeña amplitud, tiene tres componentes. La primera es la ley de Newton para el punto medio de la cuerda vocal, escrita en términos de su desplazamiento x respecto de la posición prefonatoria:

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x - by + p_g,\end{aligned}\tag{3.1}$$

donde ϵ es el coeficiente de restitución elástica de la cuerda, b es el coeficiente de disipación lineal, y p_g es la presión glotal (intercordal), promediada espacialmente en toda el área de la cuerda vocal. Todos los parámetros están definidos por unidad de masa de la cuerda.

La segunda componente del modelo es la implementación de la hipótesis de la onda viajera en la superficie de la cuerda vocal. Esto es, una imposición sobre la cinemática de las partes de la cuerda vocal. Si las áreas de entrada y salida de la glotis son a_1 y a_2 , respectivamente, se impone que

$$a_1 = 2h(x_{01} + x + \tau y)\tag{3.2}$$

$$a_2 = 2h(x_{02} + x - \tau y),\tag{3.3}$$

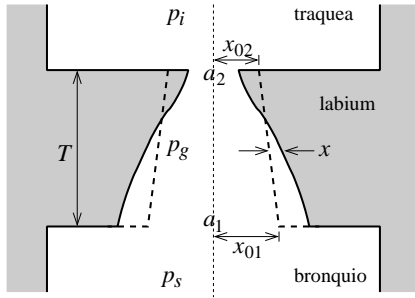


Figura 3.4: Sección frontal esquemática de la siringe, mostrando las definiciones para el modelo flameante. La variable x es el desplazamiento del punto medio del labium respecto de la posición prefonatoria. Línea discontinua: posición prefonatoria. T : espesor del labium; a_1 , a_2 : áreas de entrada y salida de la válvula vocal.

donde h es el largo antero-posterior de la glotis, x_{01} y x_{02} son las posiciones prefonatorias de los bordes inferior y superior de la cuerda, respectivamente, y $\tau = T/(2c_m)$ es el tiempo que le lleva a la onda en la mucosa atravesar la mitad del espesor de la cuerda a una velocidad c_m (ver Figura 3.4).

La tercera y última componente es una relación entre los perfiles geométricos y la presión intraglotal p_g . Titze asume que la relación entre presión y perfil geométrico en configuraciones estáticas es aplicable a flujos oscilantes [Titze 1988], y escribe

$$p_g = p_s - \left(\frac{a_2}{a_1}\right)(p_s - p_i), \quad (3.4)$$

donde p_s es la presión subglotal y p_i es la presión a la entrada del tracto vocal ($p_i = 0$ significa válvula vocal aislada, desacoplada del tracto).

El movimiento de los bordes superior e inferior de la cuerda vocal está acoplado a la oscilación del punto medio x a través de las ecuaciones (3.2) y (3.3), de forma tal que las cuerdas presentan un perfil convergente cuando se separan, y un perfil divergente cuando se aproximan. Esto hace que la fuerza ejercida sobre la cuerda vocal por la presión intraglotal p_g (ecuación 3.4) sea asimétrica en un ciclo de la oscilación, y por lo tanto permite una transferencia neta de energía del flujo de aire al tejido.

Este modelo fue concebido para estudiar las condiciones del umbral de oscilación, típicamente interpretado como el valor de presión subglotal p_s^* necesario para iniciar la oscilación, en función de las características geométricas y viscoelásticas del tejido. Más allá del umbral la amplitud de las oscilaciones diverge, pues el sistema no posee los mecanismos de saturación para modelar las colisiones de los tejidos móviles entre sí y contra las paredes. En lo que sigue mostraremos que es posible modificar el modelo flameante de Titze para utilizarlo más allá del umbral de oscilación y estudiar oscilaciones de amplitud finita.

3.3.4 Modelo extendido

El modelo flameante de Titze no fue diseñado para trabajar más allá del umbral de oscilación. Sin embargo, puede ser extendido con un término no lineal apropiado que permite obtener oscilaciones de amplitud finita. ¿Qué término no lineal es apropiado?

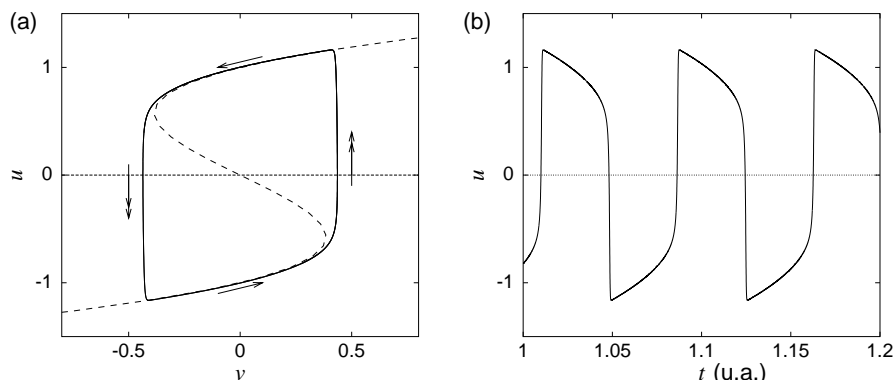


Figura 3.5: Oscilador de relajación de van der Pol con $\epsilon = 0.01$ (ecuaciones 3.5 y 3.6). (a) Espacio de fases. Línea continua: ciclo límite. Línea discontinua: nulclinas del sistema (la curva cúbica representa $\dot{u} = 0$; la recta horizontal representa $\dot{v} = 0$). (b) Serie temporal de u correspondiente al ciclo límite mostrado. Las dos escalas de tiempo son evidentes: lentas evoluciones cerca de los estados “arriba” y “abajo”, y rápidas transiciones entre estados. El sesgo característico de las oscilaciones de las cuerdas vocales humanas está exagerado aquí por el pequeño valor de ϵ .

Oscilaciones de relajación

La respuesta está en uno de los modelos paradigmáticos de la Dinámica No Lineal: el oscilador de relajación, concebido para modelar saltos entre dos estados (abierto-cerrado, tensión-relajación, etc.) por medio de un campo vector continuo y derivable [Guckenheimer and Holmes 1996, Strogatz 1994]. La forma más simple del oscilador de relajación fue dada por van der Pol:

$$\dot{u} = v - u^3 + u \quad (3.5)$$

$$\dot{v} = -\epsilon u, \quad (3.6)$$

donde ϵ es un parámetro chico. El ciclo límite en el espacio de fases se ve en la Figura 3.5a. Lentas excursiones a lo largo de la nulclina cúbica son seguidas por rápidas transiciones entre las ramas; la existencia de estos dos tiempos característicos muy diferentes es evidente en la serie temporal, que muestra el sesgo característico de las oscilaciones glotales (Figura 3.5b).

Reescribamos ahora el oscilador de van der Pol (ecuaciones 3.5 y 3.6) en la forma estándar $\dot{x} = y$, $\dot{y} = f(x, y)$. Nuestro objetivo es identificar qué término es necesario agregar en el modelo flameante de Titze (ecuación 3.1) para obtener oscilaciones de relajación. Para esto, identifiquemos la variable x en el modelo flameante con la variable u en el oscilador de relajación: $x \equiv u$. Su derivada temporal es $\dot{x} = \dot{u} = v - u^3 + u$. Llamamos $y \equiv \dot{x}$ y calculamos su derivada temporal: $\dot{y} = \dot{v} - 3u^2\dot{u} + \dot{u}$. Con esto, nos queda

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\epsilon x - 3x^2y + y. \end{aligned} \quad (3.7)$$

Notemos que estas dos ecuaciones diferenciales acopladas de primer orden son equivalentes a una sola ecuación de segundo orden: $\ddot{x} = -\epsilon x - 3x^2\dot{x} + \dot{x}$, que es la ley de Newton para un sistema puntual cuya posición es x . La interpretación de

sus términos es clara: para que el sistema presente oscilaciones de relajación es necesario tener una fuerza de restitución elástica $-\epsilon x$, una fuerza $+y$ en fase con la velocidad que aporta energía, y una disipación no lineal $-3x^2y$ que se hace importante cuando la variable x se aparta mucho del equilibrio. Este término es el encargado de simular la alta disipación que sufren los tejidos móviles cuando chocan contra las paredes de la válvula al final de la fase de apertura, o cuando chocan entre sí al terminar la fase de cierre.

Modelo extendido

Queda claro entonces que la ley de Newton del modelo flameante de Titze (ecuación 3.1) necesita el agregado de un término disipativo no lineal $-x^2y$ para mostrar oscilaciones de relajación. Con esto en mente, modificamos el modelo de Titze y escribimos

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x - by - cx^2y - f_0 + p_g,\end{aligned}\tag{3.8}$$

donde agregamos, además, un término de fuerza externa f_0 . La presión p_g y las áreas a_1 y a_2 están dadas, como antes, por las ecuaciones (3.4), (3.2) y (3.3), respectivamente. Este modelo fue el primero que utilizamos para estudiar la oscilación de la válvula vocal (ver Sección 5.1.4).

Como fue discutido más arriba en esta sección, el modelo original de Titze logra la transferencia de energía a los tejidos móviles gracias a la hipótesis de un movimiento flameante. Esta hipótesis está volcada en las expresiones para p_g , a_1 y a_2 , que conforman una fuerza altamente no lineal. Sin embargo, en la discusión del oscilador de relajación quedó en evidencia que para lograr oscilaciones autosostenidas alcanza con tener una fuerza proporcional a la velocidad. Para identificar el mecanismo de pérdida de estabilidad que lleva a las oscilaciones autosostenidas en el modelo de Titze, hacemos un desarrollo a primer orden en serie de Taylor del factor geométrico a_2/a_1 en la ecuación (3.4) alrededor del punto fijo $(x, y) = (\bar{x}, 0)$. Obtenemos

$$\frac{a_2}{a_1} \sim A - Dy + E(x - \bar{x}),\tag{3.9}$$

donde

$$\begin{aligned}A &= \bar{a}_2/\bar{a}_1 \\ D &= 2h\tau(\bar{a}_1 + \bar{a}_2)/\bar{a}_1^2 \\ E &= 2h(\bar{a}_1 - \bar{a}_2)/\bar{a}_1^2,\end{aligned}$$

y $\bar{a}_{1,2} = 2h(x_{01,02} + \bar{x})$. Notar que un perfil prefonatorio rectangular nos da $E = 0$ y $A = 1$, sin mayor perjuicio de la dinámica; en consecuencia elegimos $x_{01} = x_{02} \equiv x_0$ por simplicidad. Con esto, la expresión para p_g (ecuación 3.4) queda

$$p_g \sim p_s + (Dy - 1)(p_s - p_i),\tag{3.10}$$

con $D = 2\tau/(x_0 + \bar{x})$. Como era de esperarse, el desarrollo de Taylor del modelo flameante devela un término un fase con la velocidad. Esto se ve claramente en el caso de válvula desacoplada de tracto ($p_i = 0$), donde queda directamente $p_g = Dp_s y$.

En esta sección hemos discutido las condiciones para tener oscilaciones autosostenidas en la válvula vocal, y presentamos el modelo flameante de Titze. En base a este modelo, desarrollamos nuestro modelo extendido (ecuaciones 3.8 y 3.10), que resumimos ahora:

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x - by - cx^2y - f_0 + p_g\end{aligned}\tag{3.11}$$

$$p_g = p_s + (Dy - 1)(p_s - p_i)\tag{3.12}$$

$$p_i = p_i(x, y),\tag{3.13}$$

donde, por generalidad, permitimos que la presión p_i a la entrada del tracto vocal sea una función de las variables del sistema. El caso desacoplado ($p_i = 0$) se estudia a continuación.

Diagrama de bifurcaciones

El comportamiento del modelo extendido que desarrollamos en esta sección se resume en la Figura 3.6 para el caso desacoplado. El caso desacoplado es imponer $p_i = 0$, con lo cual el sistema queda simplemente

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x + py - cx^2y - f_0,\end{aligned}\tag{3.14}$$

donde definimos por comodidad el parámetro $p \equiv Dp_s - b$ para agrupar todos los términos con y lineales. Notemos que en este caso nuestro sistema tiene la forma estándar del oscilador de relajación de van der Pol. La aparición de oscilaciones se da cuando el punto fijo pierde estabilidad en una bifurcación de Hopf, pasando de atractor a repulsor. En otras palabras, la condición para una bifurcación de Hopf en este sistema sale de linealizar el sistema y pedir que la parte real de los autovalores cambie de signo, lo que ocurre cuando

$$p = c(f_0/\epsilon)^2.\tag{3.15}$$

En la Figura 3.6a se muestra el diagrama de bifurcaciones en el espacio de parámetros (p, ϵ) para el caso $f_0 = 0$. Los parámetros p y ϵ tienen papeles bien diferenciados: el nacimiento de las oscilaciones está gobernado por p , mientras que la frecuencia de las oscilaciones está dada por el valor de ϵ . Además, el parámetro p también determina el contenido espectral de la señal (no mostrado): mayor p significa mayor amplitud y más armónicos.

El diagrama de bifurcaciones en el espacio (p, f_0) se muestra en la Figura 3.6b. Notemos en este gráfico que el nacimiento de las oscilaciones también puede gobernarse con f_0 , como lo dice la condición de Hopf (ecuación 3.15). Esto está relacionado con la expresión para el punto fijo \bar{x} del sistema, dado por $\bar{x} = -f_0/\epsilon$; un valor grande de f_0 lleva la posición del punto fijo hacia la izquierda. Teniendo en cuenta que nuestro modelo asume simetría entre ambos labia, esto significa que un incremento en el módulo de f_0 acerca los labia uno hacia el otro (si $f_0 > 0$) hasta cerrar la siringe, o los aleja (si $f_0 < 0$) hasta abrirla completamente, a tal punto que impide la oscilación.

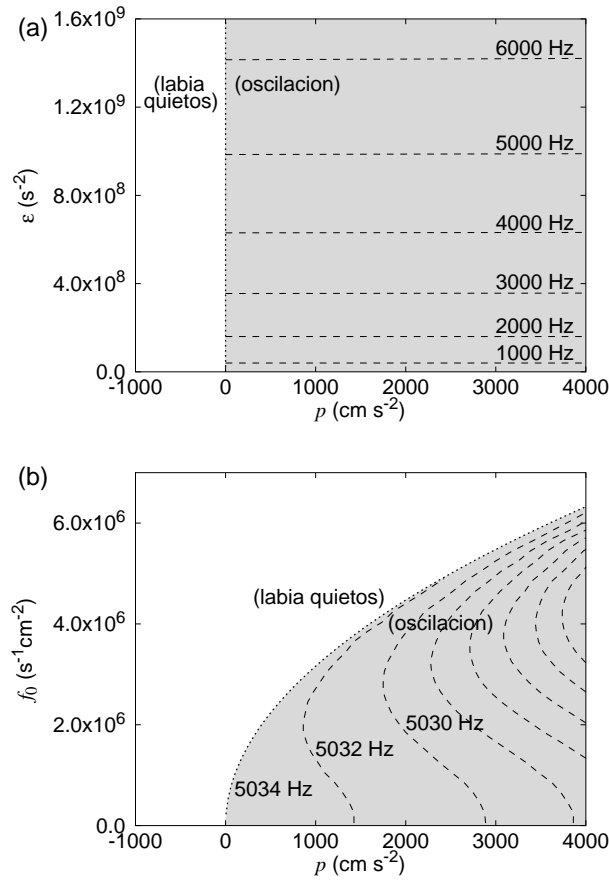


Figura 3.6: Modelo extendido, caso desacoplado (ecuación 3.14, con $p_i = 0$). (a) Diagrama de bifurcaciones en el espacio (p, ϵ) , con $f_0 = 0$. El sistema presenta oscilaciones para $p > 0$ (zona grisada). Las líneas discontinuas (prácticamente horizontales en esta escala) son los contornos de isofrecuencia. Se ve que en estos rangos los parámetros p y ϵ tienen papeles bien diferenciados: el parámetro p está asociado al nacimiento de las oscilaciones (y a la cantidad de armónicos en el espectro de la señal; no mostrado aquí), mientras que el parámetro ϵ gobierna la frecuencia de las oscilaciones. (b) Diagrama de bifurcaciones en el espacio (p, f_0) , con $\epsilon = 10^9 \text{ s}^{-2}$. El parámetro f_0 también gobierna el nacimiento de las oscilaciones. Un valor grande de f_0 (en módulo) impide la oscilación, pues lleva al punto fijo muy a la izquierda o muy a la derecha; se interpreta que la siringe está completamente cerrada o completamente abierta, respectivamente. En todo este capítulo y en el siguiente es $c = 10^8 \text{ cm}^{-2} \text{ s}^{-1}$.

Relación entre parámetros dinámicos y fisiológicos

El modelo extendido, aunque simple, permite una interpretación directa de sus parámetros en términos de las mediciones fisiológicas reportadas experimentalmente y descritas en la Sección 2.2.1.

La identificación de nuestro parámetro p_s con la presión bronquial es clara: es necesaria una presión bronquial positiva (mayor que la atmosférica) para crear un flujo de aire [Goller and Suthers 1996a], y para tener oscilaciones debe ser tal que la energía suministrada sobrepase las pérdidas por fricción. Esto es,

$Dp_s > b$ para que $p > 0$.

Con respecto al parámetro ϵ , hemos dicho que la actividad del músculo vS está directamente correlacionada con la frecuencia de la vocalización [Goller and Suthers 1995, Goller and Suthers 1996b], y que no cumple un papel en el control del flujo de aire. Es natural por lo tanto asociar la actividad en el vS con el parámetro ϵ de nuestro modelo.

El parámetro f_0 , por último, está relacionado con la actividad de los músculos aductores (dS y dTB) y abductores (vTB y TL) [Goller and Suthers 1996a]. Hemos visto que valores grandes positivos de f_0 provocan el acercamiento de las posiciones de equilibrio de los labia, a tal punto de cerrar la siringe e impedir la oscilación. Recíprocamente, valores grandes negativos de f_0 impiden la oscilación al separar los labia en exceso. El parámetro f_0 , por lo tanto, se asocia a la actividad neta del grupo aductores-abductores: f_0 toma valores positivos cuando la actividad de los dS y dTB es mayor que la de sus antagonistas vTB y TL, y valores negativos en caso contrario.

3.4 Respaldo experimental

El modelo extendido desarrollado en la sección anterior fue validado experimentalmente. Es natural que el comportamiento oscilatorio de los labia sea modelable en términos de un oscilador simple, por lo menos en alguna aproximación. Sin embargo, la interpretación de los parámetros del modelo en términos de datos fisiológicos no es trivial, y requiere validación experimental [Mindlin et al. 2003].

Mindlin y colaboradores [Mindlin et al. 2003] utilizaron datos fisiológicos reales para integrar el modelo extendido (ecuaciones 3.14) y así evaluar la capacidad del modelo para reproducir el comportamiento de los labia. Para empezar, grabaron vocalizaciones de cardenales (*Cardinalis cardinalis*) al mismo tiempo que registraban la presión en el saco aéreo P y la actividad del músculo ventral vS (que gobierna la frecuencia del sonido) y del músculo dorsal dS (involucrado en el cierre activo de la siringe). La serie temporal de P fue suavizada, y las de vS y dS fueron rectificadas y suavizadas. Las curvas envolventes resultantes fueron utilizadas para integrar el modelo, después de una simple transformación lineal ($P \rightarrow p$, $vS \rightarrow \epsilon$, $dS \rightarrow f_0$) para ajustar escalas.

El resultado se muestra en la Figura 3.7. Las sílabas superiores fueron emitidas por el ave, mientras que las inferiores fueron generadas por el modelo utilizando los datos fisiológicos de las sílabas reales. La similitud es notable, y otorga confianza a la interpretación de los parámetros de nuestro modelo [Mindlin et al. 2003].

3.4.1 Una predicción

El modelo extendido nos ofrece una predicción concreta referida a la actividad muscular al comienzo y al final de una sílaba.

El parámetro f_0 nos permite hacer una predicción sobre la actividad de los músculos que representa, los aductores dS y dTB. Recordemos el diagrama de bifurcaciones de la Figura 3.6b. La región gris indica la zona donde ocurre la oscilación de los labia; incrementar el valor de f_0 nos lleva fuera de la zona de

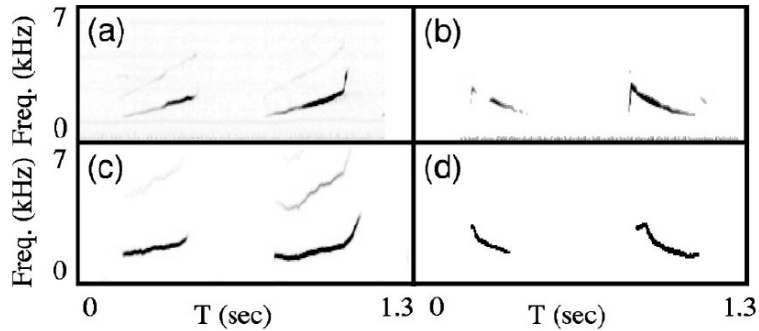


Figura 3.7: (a,b) Sonogramas de fragmentos de canto de cardenal (*Cardinalis cardinalis*). Estos sonogramas fueron registrados junto con la presión del saco aéreo y la actividad electromiográfica de los músculos vS y dS. (c,d) Sonogramas de cantos sintéticos, generados por el modelo utilizando los datos fisiológicos (presión, vS, dS) registrados en (a,b). Extraído de [Mindlin et al. 2003].

oscilación. Esto significa que los labia son llevados uno contra el otro activamente, cerrando la siringe e impidiendo la oscilación. Cuando f_0 cae por debajo de ese valor límite comienza la oscilación.

La condición para la bifurcación de Hopf (ecuación 3.15) nos dice que el valor mínimo de f_0 necesario para impedir la oscilación depende de ϵ . Esto es

$$f_0 = \epsilon \sqrt{p/c}, \quad (3.16)$$

lo que significa que el valor de f_0 necesario para impedir la oscilación es mayor si ϵ es mayor. Esto nos permite predecir que la actividad de los músculos aductores dS y dTB podría ser menor (mayor) al comienzo de una sílaba en comparación con el final, si la sílaba es de frecuencia creciente (decreciente), y si la presión se mantiene más o menos constante durante la vocalización [Mindlin et al. 2003].

3.4.2 Un experimento

Esta predicción y su verificación experimental fueron llevadas a cabo por Mindlin y sus colaboradores [Mindlin et al. 2003]. La Figura 3.8a muestra un ejemplo de series temporales fisiológicas suavizadas como las utilizadas para validar el modelo al comienzo de esta sección. Corresponden a una sílaba de frecuencia descendente, y se observa que la actividad del dS es mayor al comienzo que al final.

La predicción se verificó en ésta y todas las sílabas analizadas. La Figura 3.8b resume los resultados. El cociente entre la actividad máxima del dS al final y al comienzo de cada sílaba es mayor que 1 para todas las sílabas de frecuencia ascendente, y menor que 1 para las descendentes.

3.5 El tracto vocal

El tracto vocal es la parte del aparato vocal que va desde la fuente sonora hasta donde el sonido es emitido al exterior. En aves, va desde la siringe hasta el pico y está formado por la tráquea, la laringe, la orofaringe y el pico. En humanos,

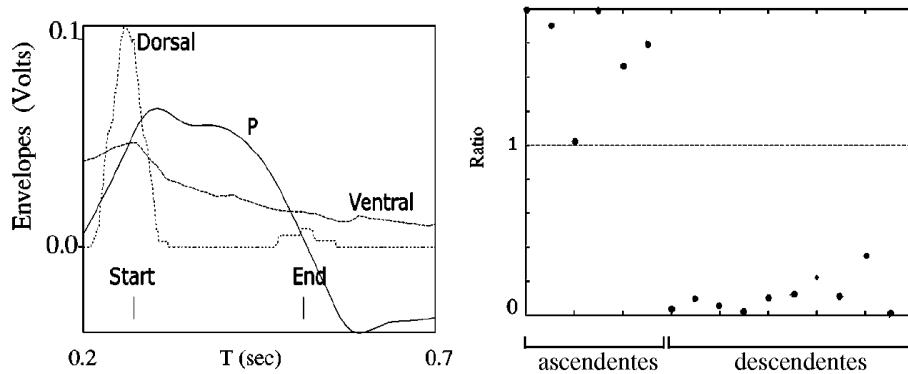


Figura 3.8: (Izquierda) Datos fisiológicos de presión en el saco aéreo (P) y actividad en los músculos que gobiernan la frecuencia (“Ventral”: vS) y el flujo de aire (“Dorsal”: dS), para una sílaba de frecuencia descendente. Se aprecia una mayor actividad del músculo dorsal (aductor) al comienzo que al final de la sílaba. (Derecha) Cociente entre la actividad máxima del músculo dS al final y al comienzo de cada sílaba. Todas las sílabas estudiadas de frecuencia ascendente tienen un cociente mayor a 1, mientras que todas las descendentes tienen un cociente menor a 1. Extraído de [Mindlin et al. 2003].

va desde la laringe hasta los labios y comprende la parte superior de la laringe, la faringe y la boca (en ambos casos el tracto vocal tiene una ramificación: la cavidad nasal). El sonido emitido por un ave lleva impresa la huella del tracto vocal, pues la forma del tracto (que puede ser alterada por la posición de la lengua, los labios o el pico, etc.) modifica el sonido emitido y puede hasta influir en las condiciones en que el sonido es producido en la válvula vocal.

3.5.1 El tracto vocal como filtro

En muchas especies de aves, y en habla humana normal, el mecanismo de fonación involucra al tracto vocal como un “filtro”. En este aspecto el tracto se comporta como una cavidad resonante pasiva que tiene ciertas frecuencias naturales: en primera aproximación se lo considera un tubo uniforme abierto en un extremo y cerrado en el otro. Esta cavidad resonante es excitada por la señal que inyecta la válvula vocal, y acentúa o atenúa diferentes componentes armónicas de la señal original de acuerdo a dónde caen sus resonancias (Figura 3.9).

El comportamiento del tracto vocal como acentuador o atenuador de frecuencias se resume en su *función de transferencia* $G(\omega)$; esto es, la respuesta del filtro ante diferentes frecuencias excitatorias. Las frecuencias de resonancia de la cavidad, tradicionalmente llamadas *formantes*, están indicadas por los picos de la función de transferencia. Si el filtro es excitado con frecuencias próximas a una resonancia, la amplitud de la oscilación se verá aumentada. En cambio, una frecuencia excitatoria que caiga entre picos de la función de transferencia verá disminuida su amplitud.

El efecto del tracto vocal como filtro es sorpresivamente perceptible en el habla humana normal al enunciar diferentes vocales. Las condiciones de oscilación de las cuerdas vocales prácticamente no cambian de vocal a vocal; sin embargo el tracto modifica el sonido al punto que podemos percibir la diferencia. La

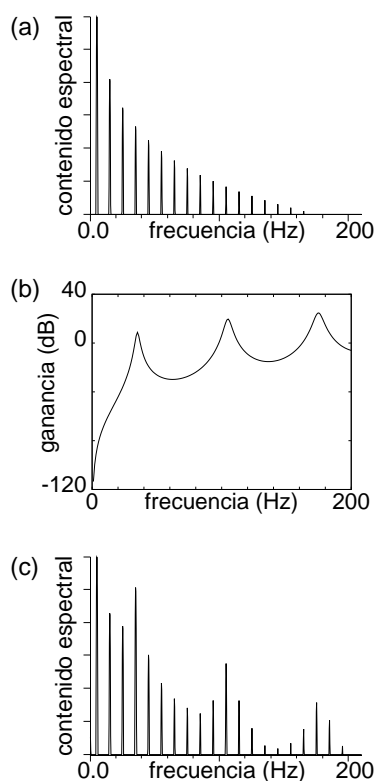


Figura 3.9: Efecto del tracto vocal cuando actúa como filtro. Una señal es inyectada en una cavidad resonante (un tubo uniforme abiertocerrado), y se mide el espectro de la señal a la salida de la cavidad. (a) Espectro de la señal original (una onda triangular). (b) Función de transferencia del filtro, o respuesta en frecuencias. Los picos son las resonancias de la cavidad. (c) Espectro de la señal resultante a la salida del filtro. La altura de los armónicos se ve modificada por el efecto del filtro.

Figura 3.10 muestra esta diferencia en forma gráfica.

En aves, ciertos experimentos dejaron en evidencia las propiedades de filtro del tracto vocal. Experimentos con heliox, una mezcla de helio y oxígeno cuya densidad es 30% la del aire normal, sugirieron que el ave modifica la configuración de su tracto vocal para acentuar determinadas frecuencias en su vocalización [Nowicki 1987]. Estos cambios en la configuración del tracto vocal se logran modificando la posición de la lengua y la apertura del pico, por ejemplo, y tendrían efecto sobre las frecuencias de resonancia del tracto (un pico más abierto acorta el largo efectivo del tracto vocal [Fletcher and Tarnopolsky 1999]). Aunque la apertura del pico tendría poca influencia sobre la frecuencia de la primera resonancia del tracto vocal [Fletcher and Tarnopolsky 1999], experimentos de manipulación de los movimientos del pico mostraron que el espectro de las vocalizaciones cambiaba en forma predecible en base a la consecuente modificación de las resonancias del tracto [Hoese et al. 2000].

El efecto de filtro en aves es patente en las vocalizaciones del guácharo (*Steatornis caripensis*), un ave con una siringe sorprendentemente asimétrica. La siringe del guácharo es bronquial (Figura 2.1c), y las longitudes de los bronquios entre cada hemisiringe y la tráquea son desiguales.¹ Las vocalizaciones sociales (no ecolocalizadoras) del guácharo tienen muchas componentes armónicas y presentan habitualmente 3 formantes. Suthers encontró que la frecuencia de las

¹El guácharo es sorprendente también por otros motivos: vive en cuevas y emite sonidos tipo "click" para ecolocalizarse en la oscuridad.

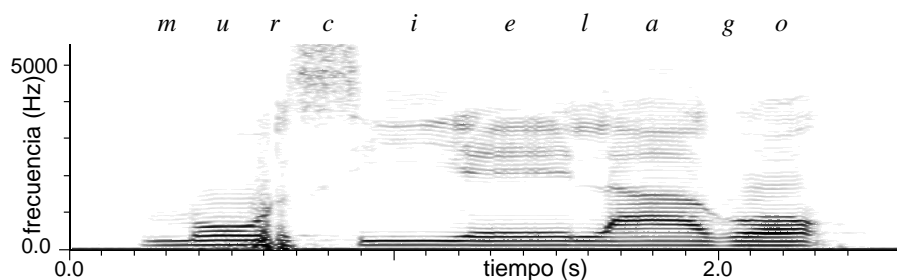


Figura 3.10: Sonograma de la palabra “murciélago”, una de las (no tan raras) palabras en español que tiene las 5 vocales sin repetir. La diferencia entre las vocales es evidente. Aunque la frecuencia fundamental es prácticamente constante, la estructura de sus armónicos es diferente. En cada caso, una reconfiguración en la posición de la lengua o los labios modificó las resonancias del tracto vocal (llamadas “formantes”), acentuando diferentes grupos de armónicos.

segunda y tercera formantes era cercana a la frecuencia de resonancia estimada para los bronquios izquierdo y derecho, respectivamente (como si fueran tubos con un extremo abierto), mientras que la primera formante podría corresponder a la frecuencia de resonancia de la tráquea (como un tubo abierto en ambos extremos) [Suthers 1994, Suthers 2001].

Separación fuente-filtro

El mecanismo recién descrito para la interpretación del papel del tracto vocal se basa en la hipótesis de *separación fuente-filtro*, considerada ampliamente en el estudio de la voz humana [Fant 1960, Rothenberg 1986]. Esta hipótesis asume que la oscilación de las cuerdas vocales es independiente de la dinámica de la presión acústica en el tracto. Es decir, considera que el sistema vocal se forma con una fuente de sonido activa no lineal (la válvula vocal) que excita un filtro pasivo lineal (el tracto), y que la propagación del sonido en el tracto no afecta la dinámica de la fuente salvo en forma pasiva.

La validez de esta hipótesis en el canto de aves fue objeto de una larga discusión. La separación fuente-filtro fue uno de los mecanismos propuestos para explicar los sonidos puros en las vocalizaciones de algunas aves (carentes de armónicos, erróneamente llamadas “silbidos”). Resultados experimentales indirectos otorgaron evidencia tanto a favor de este mecanismo [Nowicki 1987] como en contra [Ballintijn and ten Cate 1998], hasta que fue demostrado recientemente en palomas [Beckers et al. 2003a]. Otros trabajos sugieren que este mecanismo también está presente en loros [Brittan-Powell et al. 1997].

3.5.2 Cómo filtrar

El efecto de un filtro G sobre una señal s puede tenerse en cuenta de dos maneras: aplicar el filtro en el *dominio espectral*, o en el *dominio temporal*. La primera manera es de algún modo indirecta: es necesario conocer el contenido espectral de la señal $S(\omega)$ (esto es, la transformada de Fourier de la serie temporal $s(t)$) y la función de transferencia del filtro $G(\omega)$ en función de la frecuencia. El contenido espectral de la señal a la salida del filtro $P(\omega)$ está dado simplemente

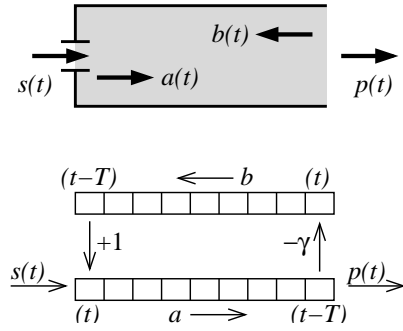


Figura 3.11: Filtrado de una señal acústica en el dominio temporal. (Arriba) El tracto vocal es considerado como un tubo uniforme. La válvula vocal inyecta la señal $s(t)$; el sonido emitido es $p(t)$. Dentro del tubo, la onda de presión que viaja hacia adelante es $a(t)$, la onda de presión hacia atrás es $b(t)$. (Abajo) Esquema del algoritmo utilizado para implementar las ecuaciones (3.18) y (3.19). Es necesario llevar cuenta de la historia de la onda de presión en el tracto, lo que se realiza con los vectores a y b .

por [Arfken and Weber 1995]

$$P(\omega) = G(\omega) S(\omega). \quad (3.17)$$

La serie temporal a la salida del filtro $p(t)$ se obtiene de antitransformar $P(\omega)$.

La segunda manera de filtrar es directa, en el sentido de que se trata de emular en forma más o menos realista el comportamiento de la señal en su pasaje a través del filtro. Por simplicidad, consideramos el tracto vocal como un tubo uniforme, cerrado en el extremo que corresponde a la válvula vocal y abierto en el otro. Dentro del tracto descomponemos la presión en dos ondas viajeras con sentidos opuestos, a hacia adelante y b hacia atrás. La onda hacia adelante a se forma con el aporte de la válvula vocal que inyecta la señal acústica s más la onda hacia atrás b proveniente de la reflexión en el extremo del tubo. Una forma de escribir esto es

$$a(t) = s(t) + b(t - T) \quad (3.18)$$

$$b(t) = -\gamma a(t - T), \quad (3.19)$$

donde γ es el coeficiente de reflexión en el extremo abierto del tubo y T es el tiempo que tarda la onda de sonido en recorrer el largo del tubo L a velocidad c ($T = L/c$) (el algoritmo para implementar este mecanismo se ilustra en la Figura 3.11). La intensidad del sonido emitido por el extremo abierto del tubo es, finalmente,

$$|p|^2 \sim (fA)^2 |a|^2, \quad (3.20)$$

donde f es la frecuencia de la onda emitida y A es el área transversal del tubo [Kinsler et al. 1982] (puede ocurrir un desfase entre p y a dependiente de la frecuencia).

3.5.3 Cuando el tracto vocal es más que un filtro

Un concepto asociado a la hipótesis de separación fuente-filtro es el de *impedancia*. La impedancia del tracto como filtro lineal pasivo es el cociente (complejo) entre la presión acústica y la velocidad de partícula.² Dicho cociente está bien definido siempre que el sistema sea lineal, pues en esos casos la velocidad adquirida por el sistema tiene siempre la misma frecuencia que la fuerza aplicada

²La impedancia de un sistema lineal en general es el cociente entre la fuerza y la velocidad; en circuitos eléctricos, por ejemplo, es el cociente entre voltaje y corriente.

(en el estacionario). La impedancia es un número complejo que depende de la frecuencia; la inversa de su módulo al cuadrado es usualmente conocida como ganancia o función de transferencia del filtro.

Si la hipótesis de separación fuente-filtro se cumple, el rol del tracto es como se describe al comienzo de esta sección: un filtro pasivo lineal, cuyo efecto se resume en su impedancia. Hasta aquí llega la habilidad de un filtro: puede acentuar o atenuar frecuencias, pero *no puede hacer aparecer* frecuencias que no estén presentes en el espectro de la señal original.

El panorama se hace mucho más amplio cuando el comportamiento del tracto vocal no es simplemente el de un filtro pasivo. En ciertas condiciones puede ocurrir que no sea posible separar los efectos de fuente y filtro. En el caso del aparato vocal, estas condiciones se cumplen cuando la válvula vocal y el tracto no son independientes sino que interactúan de alguna forma. En este sentido, puede ocurrir que la presión acústica en el tracto influya sobre la dinámica de la válvula vocal; como veremos más adelante, la interacción hace que el aparato vocal emita sonidos “inesperados” y en cierto sentido mucho más ricos.

Cuando fuente y filtro interactúan, ya no es posible utilizar la impedancia para resumir el efecto del filtro: si la fuente es no lineal, el sistema completo fuente+filtro se torna no lineal, y el enfoque de impedancias deja de ser válido. Sin embargo, el método descrito más arriba para tener en cuenta el rol del tracto vocal a través de condiciones de contorno (ecuaciones 3.18 y 3.19) tiene validez aún cuando no se cumpla la separación fuente-filtro, lo que permite estudiar el sistema en el caso interactuante. Volveremos sobre este tema en el Capítulo 5.

Capítulo 4

Gestos motores para el canto

Esta tesis está dedicada a los mecanismos físicos de la producción vocal en Oscinos. Analizamos y modelamos el comportamiento del aparato vocal del ave con las herramientas y el espíritu de la Dinámica No Lineal. Un aspecto central de esta disciplina es la búsqueda de modelos paradigmáticos que tengan un comportamiento determinado. Gracias a esta característica, los resultados de nuestra investigación nos permitieron ir más allá de la física del problema para tener una intuición sobre los mecanismos neuronales periféricos de control motor necesarios para generar las vocalizaciones. En este capítulo mostramos que una gran diversidad de sílabas puede entenderse en términos de dos gestos motores básicos: la presión en los sacos aéreos y la tensión de los labia. La riqueza de sílabas crece con la incorporación de un tercer gesto motor, el control del flujo de aire.

4.1 La estructura del canto

Empezaremos por analizar la estructura del canto y sus dos características primordiales: el ritmo, dado por la respiración, y la identidad de los sonidos individuales, dada por el dibujo de su frecuencia fundamental.

4.1.1 Sílabas

El canto de las aves canoras adultas está compuesto por una sucesión de unidades sonoras elementales más o menos estereotipadas y adquiridas por imitación, las *sílabas*. El canto se construye por repetición y alternancia de estos elementos sonoros básicos.

Las características acústicas de las sílabas y su orden dentro del canto son típicos de cada especie (ver Figura 4.1). En términos generales, podemos decir que la sílaba es una vocalización continua de alrededor de 0.1 s de duración con una frecuencia fundamental de algunos miles de Hz, aunque la variedad es común. El yacutoro (*Pyroderus scutatus*) [Straneck 1990c] y el mirasol común (*Ixobrychus involucris*) [Straneck 1990b], por ejemplo, cantan con una frecuencia fundamental de aproximadamente 200 Hz, mientras que el tangará de antifaz

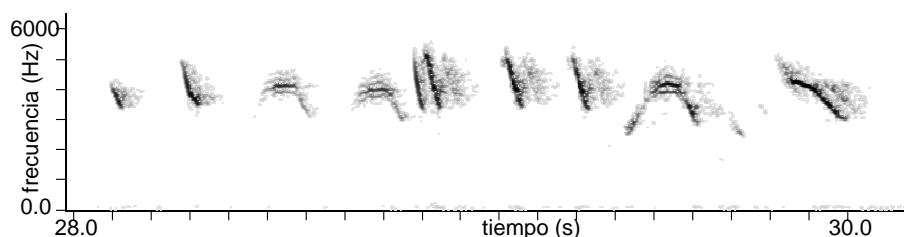


Figura 4.1: Sonograma del canto del tordo de alas amarillas (*Agelaius thilius*) [Straneck 1990a]. Cada uno de los trazos es una *sílaba*. La sílaba es una vocalización continua (no interrumpida por silencios), y es el elemento sonoro básico en la construcción del canto.

(*Pipraeidea melanonota*) [Straneck 1990c] puede llegar a los 10000 Hz. El canario doméstico (*Serinus canaria*) puede cantar sílabas tan cortas como 15 ms, en tanto que la cachirla amarillenta (*Anthus lutescens*) [Straneck 1990b] emite vocalizaciones continuas de hasta 1.5 s de duración.

Las sílabas pueden tener una frecuencia constante, aunque las más comunes son aquellas cuya frecuencia varía. La evolución de la frecuencia dentro de una sílaba puede ser creciente, decreciente, o no monótona. El contenido espectral de las sílabas también tiene gran diversidad: puede ser prácticamente nulo (un sonido “puro”, con una única frecuencia presente) o muy rico (con muchas componentes armónicas o anarmónicas), e inclusive ruidoso (una ancha banda continua de frecuencias).

4.1.2 Patrones de respiración

El canario doméstico, que tiene una frecuencia respiratoria normal de 1.5 respiraciones por segundo y un volumen tidal de apenas 0.5 ml, puede cantar sin interrupciones durante más de 30 segundos a una frecuencia de repetición de hasta 40 sílabas por segundo [Suthers 1997]. El ave logra tan altas frecuencias de repetición durante tanto tiempo gracias a *mini-inspiraciones* entre sílabas, que le permiten reaprovisionar sus sacos aéreos para seguir vocalizando y proveer oxígeno a sus pulmones (Figura 4.2). Las mini-inspiraciones fueron demostradas en 1989 [Hartley and Suthers 1989], aunque Calder ya había conjeturado su existencia en 1970, basado en evidencia indirecta [Calder 1970].

El patrón respiratorio cambia a *expiración pulsátil* si la frecuencia de repetición de sílabas es mayor que cierto límite (que para canarios es de alrededor de 30 sílabas por segundo, y para cardenales de alrededor de 16 por segundo [Suthers 1997]). Esta estrategia respiratoria no involucra mini-inspiraciones, sino que consiste en mantener una presión de saco aéreo elevada; las sílabas son generadas por interrupción periódica del flujo de aire provocada por una rápida coordinación entre los músculos aductores y abductores (Figura 4.2). Es notable que la calandria (*Mimus polyglottos*), cuando es entrenada con tutores de otra especie, utiliza en general el mismo patrón respiratorio que el tutor imitado; los límites de su desempeño vocal son evidentes al imitar trinos rápidos de canario, donde reemplaza el patrón de mini-inspiraciones utilizado por el canario por uno de expiración pulsátil, interrumpido por breves inspiraciones periódicas [Zollinger and Suthers 2004].

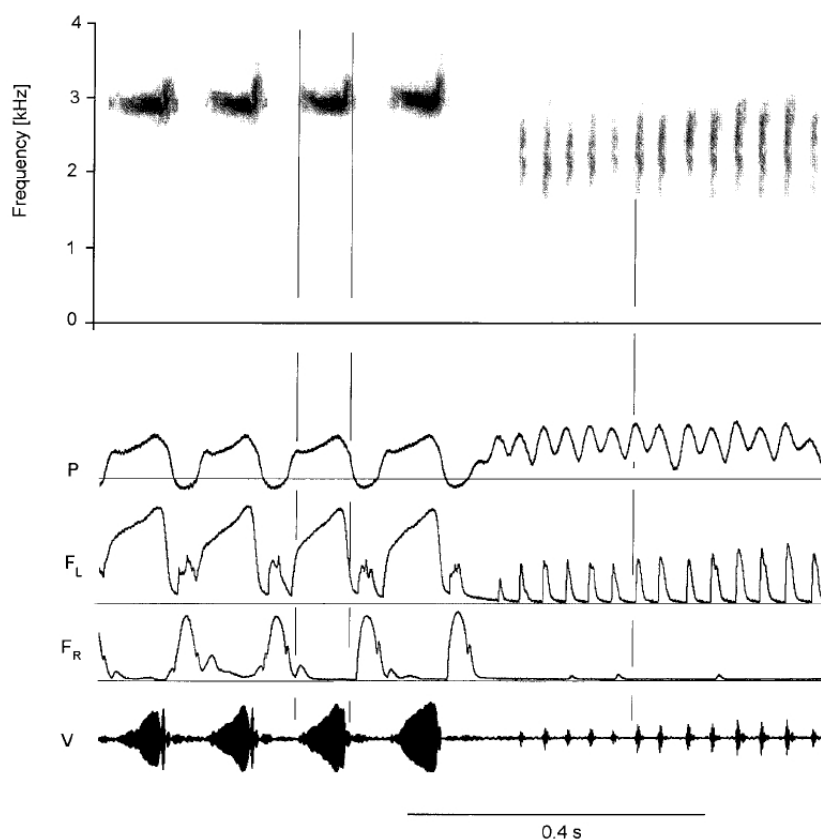


Figura 4.2: Patrones respiratorios en canarios. De arriba hacia abajo: sonograma, presión del saco aéreo (P), flujo de aire a través de los bronquios izquierdo (F_L) y derecho (F_R), y envolvente del sonido producido (V). Las sílabas de la izquierda son producidas con mini-inspiraciones: la presión del saco aéreo es negativa en los silencios entre sílabas y está acompañada por flujo de aire a través de ambos bronquios. Las sílabas de la derecha, en cambio, son producidas con expiración pulsátil: la presión del saco aéreo se mantiene siempre positiva, mientras que el flujo es interrumpido periódicamente por acción de los músculos aductores (no mostrados aquí). Extraído de [Suthers 1997].

Restricciones mecánicas y fisiológicas

El desempeño vocal del ave está en parte determinado por la capacidad para gobernar la configuración de la siringe. La existencia o no de músculos intrínsecos (Sección 2.2.3) es uno de los factores en el desempeño vocal. Además, la dificultad para producir al mismo tiempo rápidas y amplias reconfiguraciones de la siringe y el tracto con los músculos presentes puede limitar la producción vocal. Podos demostró que en gorriones pantaneros (*Melospiza georgiana*) entrenados con trinos alterados artificialmente se producían imprecisiones en la imitación debidas a restricciones motoras y no a restricciones de percepción o de memoria [Podos 1996]. En el mismo espíritu, Podos analizó los trinos de 34 especies de aves canoras y encontró un compromiso entre la frecuencia de repetición del trino y el rango de frecuencias abarcado por cada sílaba: el rango de frecuencias

es menor para frecuencias de repetición mayores [Podós 1997].

El ave enfrenta, además, otras restricciones que son más fundamentales. El patrón de respiración durante el canto es marcadamente diferente en comparación con la respiración normal del ave tranquila y sin vocalizar. Los canarios pueden cantar con mini-inspiraciones a una repetición de hasta 40 sílabas por segundo [Suthers 1997], mientras que su frecuencia respiratoria normal es de 1.5 respiraciones por segundo, como se dijo más arriba. La presión en el saco aéreo durante la fonación en el cuilacoche rojizo puede alcanzar picos con valores 50 veces mayores que los de respiración normal sin canto [Suthers et al. 1999]. Aunque estos datos indican una gran plasticidad, el aparato respiratorio no está exento de restricciones: el volumen de cada mini-inspiración en canarios corresponde al volumen de aire exhalado para producir la sílaba adyacente, y el volumen exhalado correlaciona con la duración de la sílaba [Suthers et al. 1999].

Está probado que las hembras en general son más atraídas por machos que cantan en el límite de las restricciones de la especie [Nowicki and Searcy 2004]. Se sabe por ejemplo que las hembras de canario muestran preferencia por machos que cantan sílabas con gran rango de frecuencia a una alta frecuencia de repetición [Drăgănoiu et al. 2002]. La tarea del aparato respiratorio durante el canto resulta entonces del compromiso entre un gran desempeño vocal y las restricciones, tanto mecánicas de la siringe como fisiológicas concernientes a la función respiratoria.¹

4.2 Gestos para una sílaba

En esta sección estudiamos los *gestos motores* realizados por el ave para vocalizar, es decir los ademanes musculares necesarios para emitir una sílaba. Estos gestos motores, que son más o menos estereotipados y están bien identificados experimentalmente, tienen una clara interpretación en términos de nuestros modelos.

4.2.1 Construyendo una sílaba

El modelo de fonación presentado en la Sección 3.3.4 nos da un marco para interpretar los datos fisiológicos de actividad muscular durante la vocalización y es capaz de reproducir las características principales de una variedad de vocalizaciones, como veremos a continuación. A pesar de la diversidad de sílabas existentes, veremos que muchas de ellas pueden interpretarse en términos de dos gestos motores básicos, y que una sucesión de sílabas resulta de la repetición cíclica de dichos gestos motores.

Datos fisiológicos

Las fascinantes mediciones de Goller y Suthers descritas en la Sección 2.2.1 nos permiten reconstruir la secuencia de gestos musculares utilizados por el ave para generar una sílaba (Figura 2.3).

¹Ha sido sugerido que el chotacabras europeo (*Caprimulgus europaeus*) puede cantar durante la inhalación, y de este modo lograr trinos de hasta 8 minutos de duración a una repetición promedio de 30 sílabas por segundo [Hunter 1980]. Aunque parece un hipótesis muy arriesgada, la fonación durante la inspiración fue recientemente demostrada en algunas sílabas del pinzón cebrado [Goller and Daley 2001].

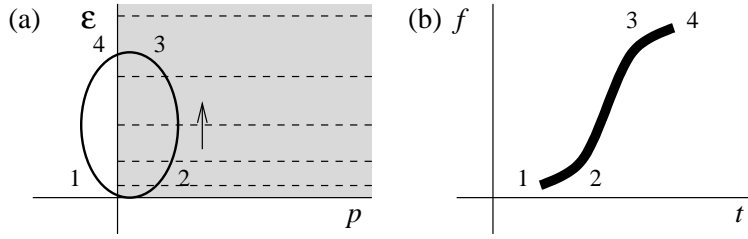


Figura 4.3: Gestos básicos para simular una sílaba ($f_0 = 0$). (a) Trayectoria simple en el espacio de parámetros (p, ϵ) que genera una sílaba de frecuencia ascendente. (b) Sonograma cualitativo resultante.

Cada sílaba comienza con una inspiración: la presión del saco aéreo cae por debajo de cero (presión atmosférica), creando un flujo de aire hacia adentro. Una vez realizada la inspiración, la presión del saco aéreo aumenta hasta que la presión intrasiríngea sobrepasa las fuerzas disipativas y comienza la fonación. A partir de ahí el protagonista es el músculo vS, cuya actividad gobierna la evolución de la frecuencia fundamental de la vocalización. En una sílaba de frecuencia creciente, por ejemplo, la actividad del vS es mayor al final que al comienzo de la sílaba. Al terminar la vocalización la actividad del vS cesa y la presión del saco aéreo cae otra vez por debajo de cero, cerrando el ciclo y dando paso a una nueva inspiración.

Una sílaba simulada

Los gestos descriptos recién tienen una clara interpretación en términos de nuestro modelo extendido (ecuaciones 3.14):

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\epsilon x + py - cx^2y, \end{aligned} \quad (4.1)$$

donde elegimos $p_i = 0$ (desacoplado del tracto) y $f_0 = 0$ por simplicidad. Podemos reconstruir la secuencia de gestos haciendo variar lentamente los parámetros del modelo, tal que reproduzcan la evolución de los parámetros fisiológicos en tiempos del orden de la duración de una sílaba. En otras palabras, buscamos una trayectoria simple en el espacio de parámetros de nuestro modelo que nos permita reproducir una sílaba.

Consideremos la trayectoria en el espacio de parámetros (p, ϵ) mostrada en la Figura 4.3a. Recordemos que el parámetro ϵ gobierna la frecuencia de la vocalización, y que el parámetro p es una transformación lineal de la presión del saco aéreo p_s .

El camino comienza en el punto 1 con una inspiración: la presión del saco aéreo es negativa (p suficientemente negativo) y los labios están relajados ($\epsilon = 0$). A continuación, el parámetro p crece para ingresar a la zona de fonación (punto 2). Como el parámetro ϵ se mantiene en valores chicos, la vocalización comienza con un sonido de frecuencia baja. A partir de ahí el valor de ϵ aumenta hasta el punto 3, lo que provoca que la vocalización tenga frecuencia ascendente. La sílaba termina cuando la trayectoria sale de la región de fonación (punto 4). El ciclo se cierra con una nueva inspiración (punto 1), listo para generar otra

sílaba. El sonograma cualitativo de la sílaba generada se muestra en la Figura 4.3b.

4.2.2 Sílabas y diferencias de fase

En base al mecanismo recién descrito para generar una sílaba de frecuencia ascendente, es fácil arriesgar que una sílaba de frecuencia descendente podría generarse con la misma trayectoria en el espacio de parámetros pero recorrida en sentido inverso. Mientras que en la sílaba de frecuencia ascendente la presión está *adelantada* con respecto a la tensión, en la sílaba de frecuencia descendente será al revés. En general, podemos pedir que los gestos cíclicos de presión y tensión estén desfasados en una cantidad arbitraria ϕ_0 ; la forma más fácil de escribir un ciclo en p y en ϵ con un desfase entre ellos es

$$\begin{aligned} p(t) &= p_0 + p_1 \cos \phi(t) \\ \epsilon(t) &= \epsilon_0 + \epsilon_1 \cos(\phi(t) + \phi_0). \end{aligned} \quad (4.2)$$

La parametrización del tiempo $\phi(t)$ nos permite controlar la velocidad a la cual el ciclo es recorrido.

En la Figura 4.4 se muestran diferentes sílabas junto con el camino en el espacio de parámetros utilizado para producirlas y la diferencia de fase correspondiente. Es notable que una gran variedad de sílabas (hacia arriba, hacia abajo, y otras no monótonas como u's y n's) son generadas simplemente al cambiar el valor del desfase ϕ_0 entre los gestos motores de presión y tensión [Gardner et al. 2001].

4.2.3 El control activo del flujo de aire

Como fue mostrado hasta aquí, la primera aproximación al control motor de la siringe debe considerar la presión del saco aéreo y la actividad del mayor músculo de la siringe, el vS, que gobierna la frecuencia del sonido emitido. Una segunda aproximación, sin embargo, debe incluir la importantísima tarea de los músculos aductores dS y dTB: permiten al ave el control activo del paso de aire a través de la siringe. El papel de estos músculos en la producción de una sílaba se describe a continuación.

El papel de los músculos aductores

El minucioso análisis de Goller y Suthers sobre el papel de los músculos siríngicos (Sección 2.2.1) nos muestra que hay detalles adicionales en la sucesión de ademanes musculares que, como veremos, le otorgan al ave una mayor libertad en el tipo de sílabas posibles.

La inspiración previa a cada sílaba se realiza con la ayuda de los músculos vTB y TL (abductores), que actúan brevemente para abrir la siringe mientras la presión del saco aéreo cae por debajo de cero, creando un flujo de aire hacia adentro. A continuación el ave utiliza los músculos expiratorios para aumentar la presión, al tiempo que cierra activamente la siringe por medio de los músculos aductores dS y dTB. Esta maniobra impide transitoriamente la fonación. Cuando la presión llega a un cierto valor, la actividad de los aductores cae; la siringe

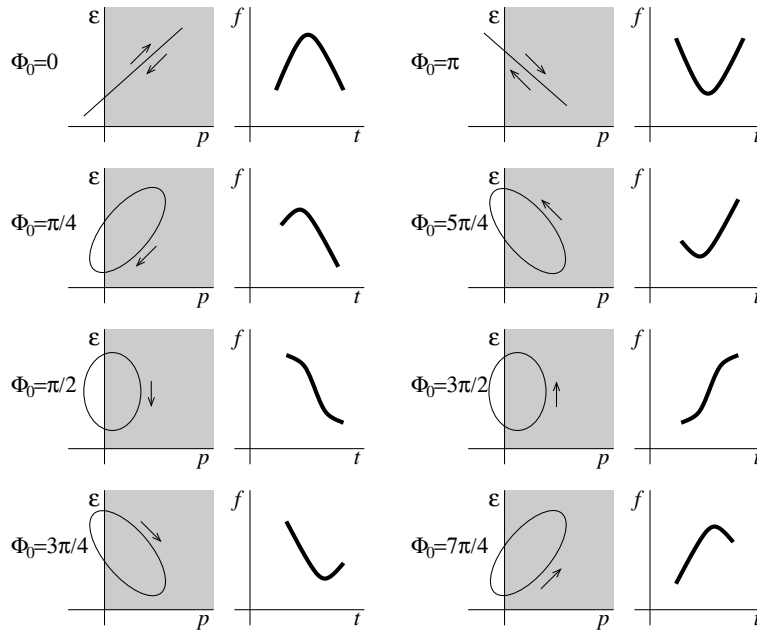


Figura 4.4: La diferencia de fase ϕ_0 entre los gestos cíclicos de presión p_s y tensión ϵ (ecuaciones 4.2) es el parámetro clave que le otorga identidad a la sílaba. La diferencia de fase ϕ_0 se incrementa de un panel al siguiente en pasos de $\pi/4$ (de arriba hacia abajo y luego de izquierda a derecha). En cada panel se muestra cualitativamente la trayectoria en el espacio de parámetros y el sonograma resultante.

se abre y comienza la fonación, a una frecuencia determinada por la actividad del vS. Para finalizar la sílaba el ave activa nuevamente los aductores durante un breve lapso, cerrando activamente la siringe para terminar la oscilación. La presión cae y el ciclo comienza otra vez.

En resumen, la actividad del grupo de músculos aductores (dS y dTB) le otorga al ave un mayor control del comienzo y final de la fonación. Sus *antagonistas*, los músculos abductores (vTB y TL), en cambio, no tienen mayor tarea que la de forzar la apertura de la siringe para permitir una rápida inspiración.²

Un ejemplo: el canto del chingolo

El canto del chingolo (*Zonotrichia capensis*), un ave muy común en toda América Central y del Sur, está formado por una introducción y un trino (Figura 4.5a). La introducción es una serie de dos a cinco sílabas ascendentes o descendentes, mientras que el trino es la repetición de una sílaba un número variable de veces. La segunda parte del canto, que presenta poca variabilidad entre individuos de una misma región, cambia consistentemente de una región a otra, dando lugar a *dialectos* regionales [Nottebohm 1969] asociados a la vegetación presente [Tubaro et al. 1993].³

²El parámetro f_0 en nuestro modelo nos permite agrupar, de ser necesario, los músculos aductores y sus antagonistas en una sola entidad.

³Es notable que el trino del chingolo es más lento en dialectos de áreas boscosas, donde la acumulación de ecos puede distorsionar una señal rápidamente modulada [Tubaro et al. 1993].

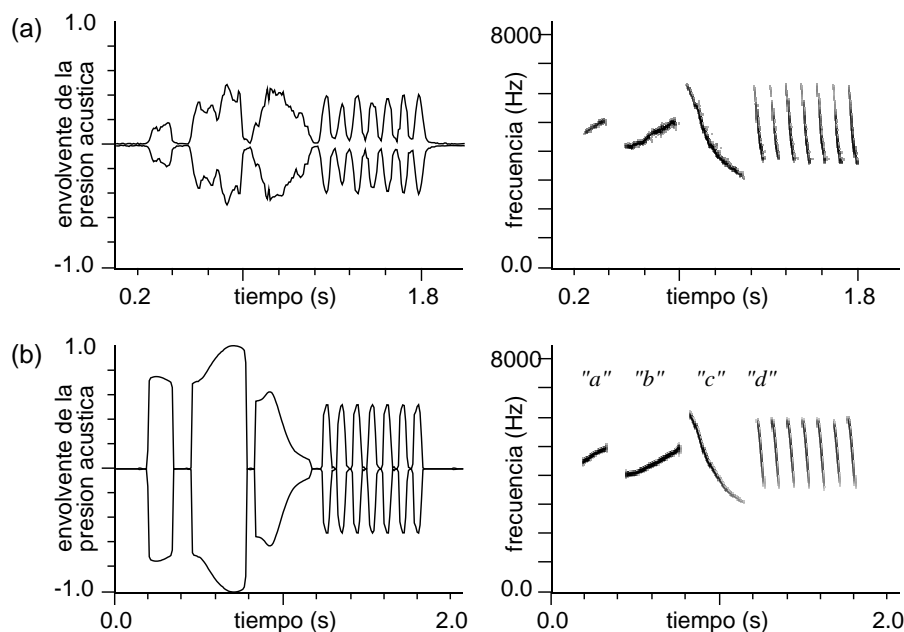


Figura 4.5: (a) Serie temporal y sonograma del canto del chingolo (*Zonotrichia capensis*), compuesto por tres sílabas introductorias y un trino con número variable de sílabas. (b) Serie temporal y sonograma de un canto sintético, obtenido al integrar el modelo (ecuaciones 4.3) a lo largo del camino mostrado en la Figura 4.7.

A pesar de que las sílabas del chingolo son diversas (sílabas ascendentes y descendentes que abarcan diferentes rangos de frecuencia), es posible elegir un camino muy simple en el espacio de parámetros de nuestro modelo para generar un canto sintético muy similar al original. Continuaremos utilizando el modelo extendido desarrollado en el Capítulo 3 (ecuaciones 3.11–3.13). Dada la simplicidad de los elementos sonoros en el canto del chingolo, consideramos que fuente y tracto están desacoplados, y por lo tanto tomamos $p_i = 0$ (Sección 3.5.3). Con esta consideración el modelo queda

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x + py - cx^2y - f_0,\end{aligned}\tag{4.3}$$

que es el mismo que utilizamos al comienzo de esta sección; la diferencia es que ahora permitimos que el parámetro f_0 (asociado a la actividad de los músculos aductores) tome valores diferentes de cero. El objetivo es encontrar una trayectoria continua en el espacio de parámetros (p, ϵ, f_0) para generar un canto sintético que reproduzca cualitativamente el canto del chingolo.

Un camino en el espacio de parámetros

Para reproducir el canto del chingolo, tomamos la idea de “gestos motores” presentada en la subsección anterior. En este sentido, utilizamos la elipse en el espacio de parámetros (p, ϵ) como el ladrillo básico para construir el canto, y le superponemos el papel del parámetro f_0 para dar cuenta de los gestos

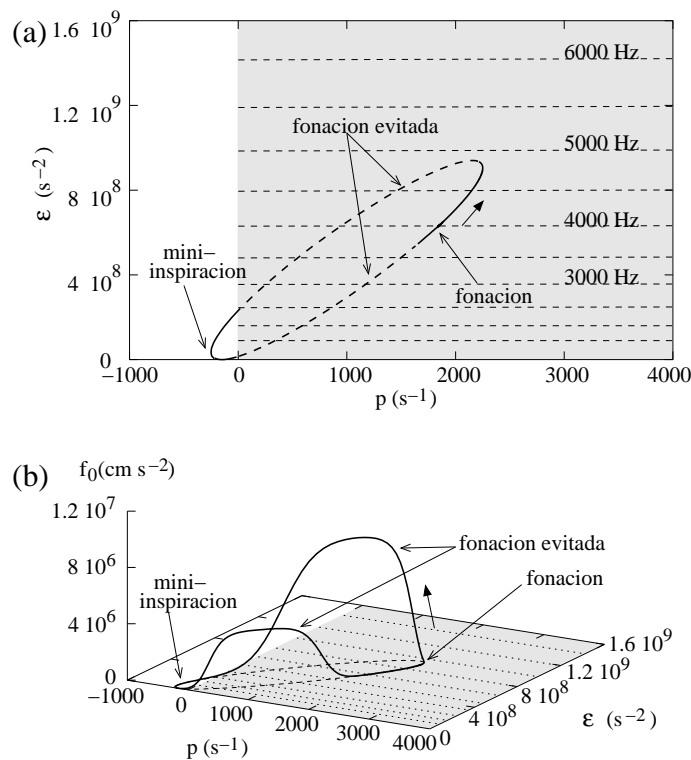


Figura 4.6: Camino en el espacio de parámetros (p, ϵ, f_0) para simular la primera sílaba del chingolo. (a) Proyección bidimensional en el espacio (p, ϵ) . Las partes de la elipse en línea discontinua corresponden a valores grandes de f_0 (siringe cerrada, no hay fonación). La fonación ocurre a lo largo de las partes de la elipse con línea continua ($f_0 = 0$) que se encuentran en la zona gris. (b) Vista 3D. La elipse se muestra en la base del gráfico. ($c = 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ en este capítulo y en el anterior).

aductores descriptos más arriba. En la Figura 4.6a se muestra la elipse elegida para reproducir la primera sílaba del canto.

La trayectoria mostrada es una forma muy simple de reproducir los gestos motores descriptos anteriormente. El ciclo comienza con un inspiración (p ligeramente negativo), con los labia relajados ($\epsilon = 0$) y la siringe abierta ($f_0 = 0$). El parámetro p crece para entrar en la zona de fonación, pero la oscilación prematura de los labios es evitada por un alto valor de f_0 (tramo de línea discontinua). Cuando p y ϵ llegan a un cierto valor, el parámetro f_0 cae a cero y comienza la fonación. La evolución de la frecuencia del sonido está dada por la evolución de ϵ (tramo de línea continua), hasta que el valor de f_0 crece abruptamente otra vez para terminar la fonación. A partir de ahí p y ϵ regresan a sus valores iniciales (tramo de línea discontinua) para eventualmente recomenzar el ciclo. La trayectoria en el espacio de parámetros tridimensional se muestra en la Figura 4.6b [Laje et al. 2002].

Ahora construyamos el canto completo. El canto sintético completo se forma con una sucesión de elipses como la descripta recién. Cada una de las sílabas del canto estará representada por una elipse apropiada; las elipses correspondientes a sílabas consecutivas deben empalmarse suavemente en la región de

mini-inspiración. Las amplitudes de las elipses en el eje horizontal están dadas por la intensidad sonora relativa de cada sílaba. Las amplitudes en el eje vertical están dadas por la frecuencia máxima de cada sílaba. La evolución del parámetro f_0 (encargado de retrasar el comienzo de la fonación o de anticipar su final) debe estar coordinada con la evolución de p y ϵ a lo largo de las elipses, como fue discutido recién.

El resultado está en la Figura 4.5b. El acuerdo con el canto original es notable. La trayectoria utilizada para el canto sintético completo y las series temporales correspondientes se muestran en la Figura 4.7 [Laje et al. 2002].

4.2.4 Sílabas con diferente patrón respiratorio

Una advertencia tiene lugar aquí. El mecanismo propuesto en esta sección para reproducir vocalizaciones de aves se basa en el análisis de cantos cuyas sílabas están separadas por mini-inspiraciones. Sin embargo, como describimos en la Sección anterior, la estrategia respiratoria de mini-inspiraciones no es la única utilizada por las aves. La estrategia de expiración pulsátil reemplaza a la primera cuando la frecuencia de repetición silábica es alta. Es de esperar que una ligera variación del mismo mecanismo sea capaz de reproducir las sílabas de expiración pulsátil mostradas en la Figura 4.2, por ejemplo, pero todavía no está demostrado. Más trabajo todavía hace falta para interpretar las sílabas de expiración pulsátil mostradas en [Goller and Suthers 1996a], que son generadas con una presión del saco aéreo prácticamente constante.

4.3 De la siringe al cerebro

Mostramos en este Capítulo que una gran variedad de vocalizaciones pueden ser interpretadas en términos de dos gestos motores básicos: presión en el saco aéreo y tensión de los labia. Una pregunta es natural: ¿dónde se generan dichos gestos motores? Estos gestos, que son principalmente gestos musculares, deben ser generados como instrucciones neurales en algún grupo de neuronas en el cerebro. ¿Cuáles son los centros neuronales encargados de enviar las instrucciones a los músculos de la fonación? Diferentes músculos son enervados por diferentes grupos de neuronas. ¿Cómo se logra la coordinación necesaria entre los distintos gestos motores?

4.3.1 El cerebro de los Oscinos

Desde la identificación de regiones del cerebro que controlan el canto en aves canoras [Nottebohm et al. 1976], el sistema del canto se convirtió en uno de los más importantes modelos para estudiar aprendizaje y control motor en vertebrados, combinando un comportamiento fácilmente registrable y un sustrato neuronal discreto [Brenowitz et al. 1997]. Esto llevó a la acumulación de una gran cantidad de información sobre las estructuras neuronales que subyacen a este comportamiento.

El canto de las aves canoras es regulado por una red discreta de centros neuronales interconectados entre sí. Estos centros o núcleos neuronales son grupos de neuronas con características similares; usualmente se destacan de los tejidos que los rodean (aún en preparaciones sin tinción) y tienen conexiones en

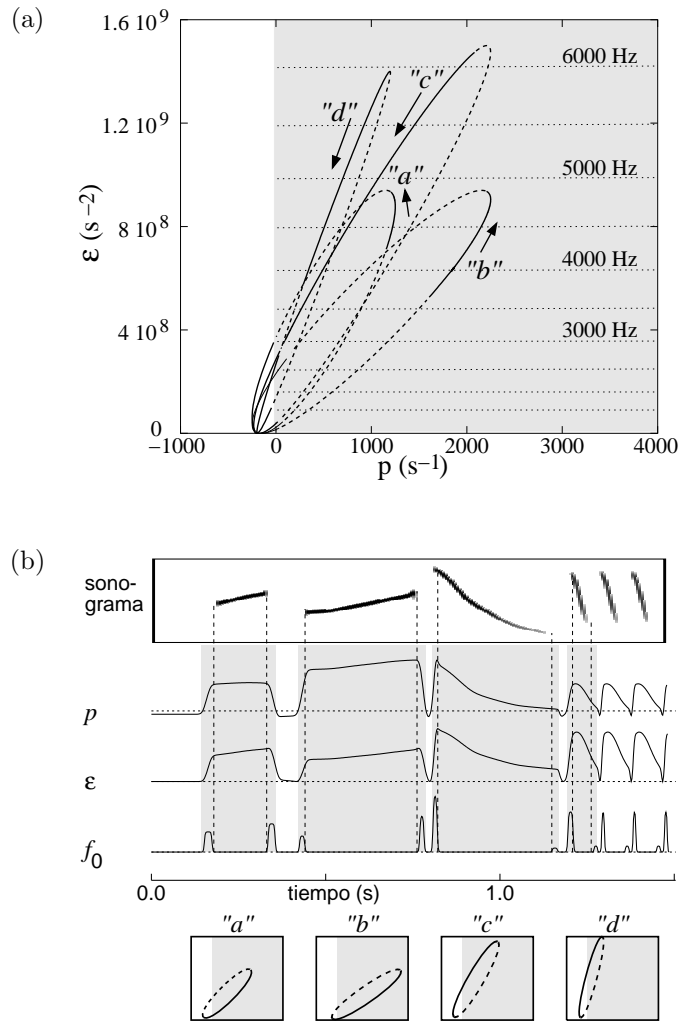


Figura 4.7: Evolución de los parámetros para generar el canto sintético mostrado en la Figura 4.5b. (a) Trayectoria continua en el espacio de parámetros. Cada elipse corresponde a una sílaba (las flechas indican sentido de recorrido). Elipses consecutivas empalman suavemente en la región de mini-inspiración. La parametrización del tiempo $\phi(t)$ no se muestra; las partes en línea continua son recorridas más lentamente que las partes en línea discontinua. (b) Series temporales de los parámetros (las elipses correspondientes se muestran debajo). Notar que cada sílaba comienza y termina con un “burst” del parámetro f_0 .

gran medida restringidas a otros núcleos del canto. Además parecen ser funcionalmente específicos al canto, por lo que forman un sistema aparentemente autocontenido que es suficiente para explicar la mayoría de los aspectos del comportamiento [Farries 2001, Farries 2004].

El corazón del sistema del canto está compuesto por 5 núcleos neuronales divididos en 2 caminos: el *camino de control motor* y el *camino del cerebro anterior* o *AFP*. A continuación describimos brevemente la estructura y función de estos caminos.

El camino de control motor

El camino de control motor, formado por los núcleos HVC y RA (ver Figura 4.8), controla la producción de canto. Estudios de lesiones dejaron establecido que sin estos núcleos el ave adopta la postura del canto pero es incapaz de vocalizar [Nottebohm et al. 1976].

El camino de control motor comienza en el núcleo HVC; este núcleo manda proyecciones al RA, que a su vez se proyecta a una variedad de núcleos motores medulares, entre los que se encuentran el nXIIts y el PAm/RAM. El nXIIts está formado por neuronas motoras que mandan sus axones a los músculos de la siringe, mientras que el PAm/RAM contiene neuronas que se activan en fase con la inspiración/expiración [Farries 2001, Brenowitz et al. 1997].

Es sumamente interesante la organización del camino de control motor. Aunque el papel individual de cada núcleo dista de ser claro, es de destacar que parece haber una cierta jerarquía en la estructura. La microestimulación del HVC interrumpe el canto, comenzando otra vez desde el principio; en cambio, la microestimulación del RA interrumpe solamente la sílaba que está siendo emitida [Vu et al. 1994].

Hay otras características de la organización del camino de control motor que merecen ser comentadas. Las instrucciones motoras para el canto, que comienzan en el HVC, se separan en partes siringea y respiratoria a nivel del RA. Este núcleo tiene subpoblaciones de neuronas excitatorias que se proyectan al nXIIts y al PAm/RAM, como ya dijimos, y tiene además interneuronas inhibitorias que establecen conexiones entre estas dos subpoblaciones [Spiro et al. 1999]. Ha sido sugerido que el RA es una especie de “caja de conexiones” para la interacción entre las instrucciones que van a la siringe y las que van a los músculos respiratorios [Mindlin and Laje 2005]. Las instrucciones siringeas y respiratorias se interrelacionan también a un nivel más bajo: se conoce que PAm/RAM envía axones a nXIIts, y se supone que las conexiones entre estos núcleos medulares pueden ayudar a la coordinación de la vocalización sin la intervención de regiones más altas del cerebro [Farries 2001].

Además de estas partes centrales, hay otras regiones del cerebro asociadas con la producción del canto cuyas funciones son menos conocidas. El núcleo NIf, por ejemplo, envía al HVC y podría transmitir instrucciones premotoras (funcionando como una especie de reloj) e información auditiva [Farries 2001, Farries 2004].

El camino del cerebro anterior (AFP)

El AFP es un conjunto de núcleos no involucrados directamente en la producción de canto: estos núcleos pueden ser lesionados en el adulto sin una degradación

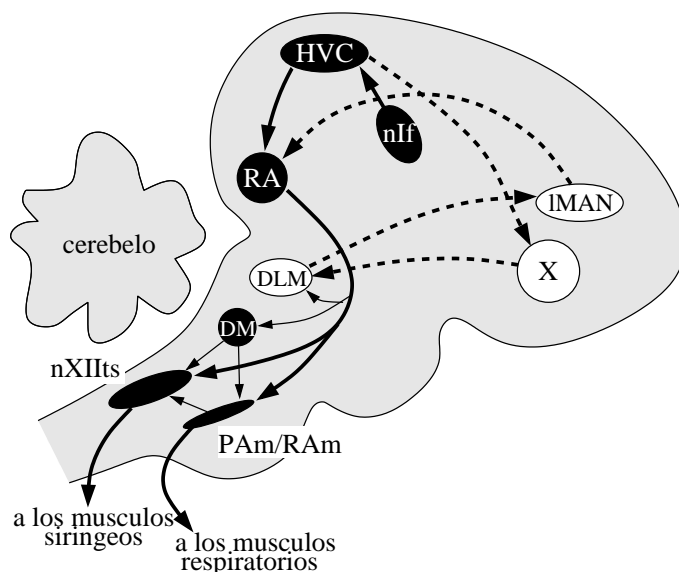


Figura 4.8: Vista sagital esquemática del cerebro de los Oscinos. Las áreas en negro son necesarias para la producción del canto adulto, mientras que las blancas son necesarias para el desarrollo del canto pero no para la producción adulta. Las flechas de línea continua forman el camino de control motor; las flechas de línea discontinua forman el camino del cerebro anterior o AFP. Modificado de [Spiro et al. 1999].

inmediata del canto. Sin embargo, son esenciales para el aprendizaje en jóvenes y para el mantenimiento del canto en adultos [Brenowitz et al. 1997, Farries 2001, Farries 2004].

El corazón del AFP está formado por tres núcleos: área X, DLM y IMAN (Figura 4.8). El área X recibe instrucciones desde el HVC y se proyecta al DLM, que a su vez se proyecta al IMAN; IMAN se proyecta sobre el RA. La función de estos núcleos durante el aprendizaje es desconocida. Se sabe que la lesión del área X en ejemplares jóvenes impide la cristalización del canto, mientras que la lesión del IMAN lleva a un canto de estructura estable pero aberrante [Brenowitz et al. 1997]. Además, el AFP estaría involucrado en el procesamiento de información auditiva referida al canto: neuronas en el AFP muestran respuestas selectivas frente a la audición de diferentes cantos [Farries 2004].

El camino de control motor y el AFP no están aislados entre sí. Como fue señalado, el HVC (en el camino de control motor) envía proyecciones al área X (en el AFP). Recíprocamente, el IMAN se proyecta sobre el RA; durante el aprendizaje, esta conexión es la encargada de enviar las señales de error o de refuerzo que ayudan a construir la arquitectura de conexiones entre HVC y RA necesaria para producir el canto adulto [Bottjer 2002].

4.3.2 Sustrato neuronal para la diferencia de fase

El canto requiere la coordinación de los grupos musculares siríngeos y respiratorios. Los patrones de activación de estos músculos son generados por los núcleos del camino de control motor, HVC y RA. ¿En qué parte del camino se establece

la coordinación entre los diferentes gestos musculares? En particular, ha sido sugerido que los patrones de actividad en el RA son el resultado de la interacción entre circuitos locales e instrucciones enviadas desde el HVc [Yu and Margoliash 1996]. Es natural esperar, por lo tanto, que la arquitectura de conectividades entre el HVc y el RA, y dentro del RA, juegue un papel preponderante en la determinación de las características acústicas de las sílabas [Abarbanel et al. 2003].

La plausibilidad de esta idea fue teóricamente demostrada por Mindlin y sus colaboradores [Abarbanel et al. 2003], con un modelo del camino de control motor que reproduce (en menor escala) las poblaciones neuronales de los núcleos HVc y RA. En dicho trabajo, las conexiones entre las neuronas simuladas del HVc y las del RA estaban representadas por los elementos (1 o 0) de una matriz de conectividades. Probaron que diferentes matrices de conectividad (es decir, diferentes conexiones entre HVc y RA) producen sílabas cualitativamente diferentes. Además, probaron que la manera de interconectar las dos subpoblaciones de neuronas dentro del RA (que envían instrucciones a los núcleos siríngeos y respiratorios, respectivamente) influye sustancialmente en la identidad de las sílabas producidas [Abarbanel et al. 2003].

Un modelo mucho más simple publicado previamente [Laje and Mindlin 2002] intentó establecer una relación entre la arquitectura de conexiones en el RA y las sílabas producidas. Dicho trabajo, sin embargo, falla en capturar aspectos de la actividad neuronal descritos posteriormente, en particular que las neuronas del HVc que se proyectan al RA no presentan actividad tónica (constante) sino esparcida (“bursts”, o grupos de potenciales de acción separados por silencios) [Fee 2002, Abarbanel et al. 2003]. Es evidente que el núcleo RA es una pieza clave para determinar la diferencia de fase entre los gestos motores de presión y tensión, pero el mecanismo dista de ser entendido en su totalidad, en particular en lo que se refiere al rol de la arquitectura interna del RA en la fase de aprendizaje [Abarbanel et al. 2003]. ¿Cuáles son los límites del control motor del canto? Más allá de las restricciones mecánicas del aparato fonador y fisiológicas del aparato respiratorio, ¿puede el ave aprender una combinación arbitraria de gestos musculares, o existen restricciones a nivel neuronal que se lo impiden?

Capítulo 5

Sonidos complejos

El mecanismo básico de fonación en aves es ahora conocido, gracias a los experimentos y la teoría mencionados en los Capítulos 3 y 4. No obstante, el aparato vocal del ave despliega una increíble capacidad para producir sonidos, más allá de los que analizamos hasta ahora: sonidos que no se pueden explicar con las herramientas descritas hasta aquí. Este capítulo trata de dichos sonidos y de algunos mecanismos físicos que podrían dar cuenta de la riqueza acústica que despliega el ave en sus vocalizaciones.

5.1 Sonidos complejos

El mecanismo básico de producción de sonido en Oscinos es muy similar al utilizado por los humanos para generar sonidos voceados en habla normal. Como la laringe en humanos, la siringe en Oscinos genera sonido cuando unas masas de tejido conectivo móviles, los labia, empiezan a oscilar espontáneamente inducidas por el flujo de aire. Esta oscilación modula a su vez el flujo de aire dando lugar a una perturbación acústica que se propaga a lo largo del tracto vocal.

Hasta hace poco tiempo, la investigación sobre la producción sonora en canto de aves estuvo principalmente dedicada a develar los mecanismos físicos responsables de la acústica de la siringe, tanto en Oscinos [Goller and Larsen 1997b], como en loros [Brittan-Powell et al. 1997] y palomas [Ballintijn and ten Cate 1998]. Una buena cantidad de preguntas básicas sobre la producción de sonido están todavía abiertas, especialmente en lo que concierne a sonidos con características complejas como saltos de frecuencia, duplicaciones de período y frecuencias subarmónicas, y fenómenos como interacción fuente-fuente, etc. La existencia de estos fenómenos comenzó a ser analizada recientemente desde la perspectiva de la Dinámica No Lineal [Fee et al. 1998, Fletcher 2000].

La ocurrencia de sonidos complejos está bien establecida en las vocalizaciones de muchos mamíferos [Wilden et al. 1998], y en particular en el caso de humanos [Wilden et al. 1998, Švec et al. 1996]. Los orígenes dinámicos de tales comportamientos distan de ser claros, sin embargo. Aunque hay acuerdo en la naturaleza no lineal de los fenómenos observados (subarmonicidades, bifonación y dinámica caótica, por ejemplo), el análisis ha estado mayormente restringido a la interpretación de estos fenómenos en términos de modelos de alta dimensión [Švec et al. 1996, Steinecke and Herzel 1995]. La incipiente investigación de

sonidos complejos en vocalizaciones de aves ha comenzado su camino dentro de este paradigma [Fee et al. 1998].

5.1.1 Instrucciones vs. mecánica

Analicemos los sonogramas mostrados en la Figura 5.1. No hemos analizado sonidos tan complicados hasta ahora. El primer sonograma (Figura 5.1a) muestra una frase del canto del zorzal de cabeza gris (*Turdus subalaris*) [Straneck 1990c]. Podemos identificar dos series de trazos que se organizan de manera complicada: los trazos que componen la “línea superior”, y los que componen la “línea inferior”. Los trazos de la línea superior no comienzan ni terminan cuando lo hacen los de la línea inferior, y tampoco están relacionados armónicamente. La evolución temporal y espectral completamente diferente de estas dos líneas sugiere que el ave está utilizando las dos fuentes en forma independiente. El sonido está formado por la superposición de las contribuciones independientes de cada fuente, y la contribución de cada fuente puede entenderse en términos de los modelos simples desarrollados en el Capítulo 3. La física del problema es tan simple como antes; la complejidad de este canto radica en la naturaleza de las instrucciones neuronales que van a cada lado de la siringe, y en la coordinación necesaria entre ellas para lograr este bellísimo “duo interno”.

Comparado con el sonograma recién descrito, el de la Figura 5.1b es engañosamente simple. Se trata de una sílaba del canto del cachilo de ceja amarilla (*Ammodramus humeralis*, también conocido como pajero común) [Straneck 1990a], un ave presente en toda Sudamérica con excepción de la Patagonia argentina y la región andina. La sílaba comienza con un trazo bien definido de frecuencia ligeramente decreciente, pero a los 0.35 s aproximadamente aparecen trazos paralelos que no son armónicos del trazo original. Un análisis minucioso de los valores de frecuencia de cada trazo nos devela detalles muy interesantes. La frecuencia del trazo original justo en la transición es de aproximadamente 3900 Hz. Podría pensarse que, por alguna razón desconocida, la frecuencia fundamental del sonido en la transición pasó de 3900 Hz a alguna frecuencia muy baja, tal que todos los trazos a la derecha de la transición sean sus armónicos. Esta hipótesis es poco probable. La frecuencia de los trazos paralelos más cercanos al trazo central de 3900 Hz es de aproximadamente 4225 Hz y 3575 Hz, respectivamente. La diferencia de frecuencia entre trazos adyacentes es de aproximadamente 325 Hz, un valor probablemente demasiado bajo para esta pequeña ave.

5.1.2 Subarmónicos

El sonograma mostrado en la Figura 5.1b tiene más cosas que decir. A nadie le sorprendería encontrar, en una vocalización humana normal, armónicos que aparecen y desaparecen del espectro de acuerdo a la vocales pronunciadas, como en el sonograma de la Figura 2.2b. Este fenómeno se explica bien con los modelos que tratamos hasta ahora: tenemos una fuente de sonido (las cuerdas vocales) con muchas componentes armónicas y un filtro pasivo lineal (el tracto vocal) que acentúa o atenúa armónicos según sea su configuración (apertura de la boca, posición de la lengua, etc.). Sin importar la aparición o desaparición de armónicos, la fuente de sonido sigue vibrando con el mismo período. Esto puede

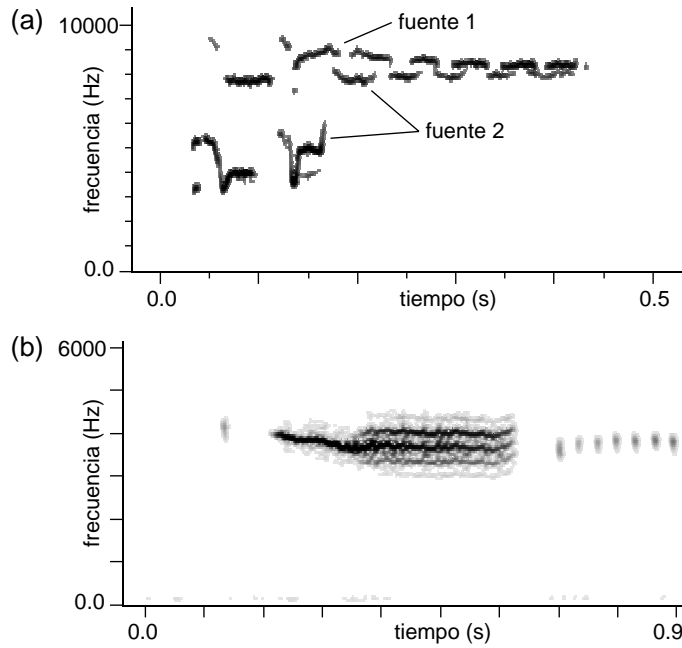


Figura 5.1: Diferentes orígenes para la complejidad en un sonograma. (a) Una frase del canto del zorzal de cabeza gris (*Turdus subalaris*). En apariencia complicado, este sonograma puede explicarse con los modelos simples de fonación desarrollados en el Capítulo 3. El ave utiliza ambas fuentes en forma independiente, y la complejidad del sonograma está dada directamente por las instrucciones neuronales. (b) Una sílaba del canto del cachilo de ceja amarilla o pajero común (*Ammodramus humeralis*). Los trazos paralelos que aparecen aproximadamente en 0.35 s son la huella de una dinámica compleja a nivel de la siringe, y no pueden explicarse en base a los modelos descritos hasta ahora.

verse al considerar la siguiente onda:

$$f(t) = A \cos(\omega t) + \epsilon \cos(2\omega t), \quad (5.1)$$

que tiene período $T_0 = 2\pi/\omega$, sin importar el valor de ϵ . Su espectro consiste en una frecuencia fundamental $F_0 = 1/T_0$ con amplitud A , más su primer armónico $F_1 = 2F_0$ con amplitud ϵ . En cambio, una onda que adquiere un *subarmónico* cambia repentinamente su periodicidad. La onda

$$f(t) = A \cos(\omega t) + \epsilon \cos\left(\frac{1}{2}\omega t\right). \quad (5.2)$$

tiene período $T_0 = 2\pi/\omega$ como antes si $\epsilon = 0$, pero pasa repentinamente a tener período *doble* $2T_0$ tan pronto como ϵ sea diferente de cero. Este tipo de huellas en un espectro no se pueden explicar con los modelos desarrollados en el Capítulo 3.

Ante la falta de evidencia experimental directa que demuestre lo contrario, el sonograma de la Figura 5.1b se interpreta en términos de una sola fuente sonora activa. Los detalles del espectro, sin embargo, nos dan pistas sobre la existencia de fenómenos físicos que están ocurriendo en la siringe de los que no podemos dar cuenta con nuestros modelos simples.

5.1.3 Paradigma de estructura intrínseca

Fenómenos complejos de una sola fuente como el mostrado en la Figura 5.1b son habitualmente interpretados en términos del comportamiento no lineal de la estructura intrínseca de la fuente sonora. Por ejemplo, un mecanismo por el cual el órgano vocal del ave agrega complejidad al sonido fue explorado en el marco del modelo de dos masas [Fee et al. 1998], un modelo tomado directamente de los trabajos sobre cuerdas vocales humanas [Ishizaka and Flanagan 1972] (ver Sección 3.1.3). El modelo mostró un comportamiento cualitativo similar al observado en siringes *in vivo* e *in vitro*, con duplicaciones de período, *mode-locking*, transiciones entre dinámica periódica y no periódica y otros comportamientos no lineales [Fee et al. 1998].

Como en el caso de fonación en humanos [Steinecke and Herzel 1995], en general se supone que las vocalizaciones complejas en aves son generadas por una dinámica no lineal intrínseca de la vibración del aparato mecánico. Este paradigma asume que el tracto vocal es un simple filtro pasivo lineal que no afecta la dinámica de la fuente; esta hipótesis, discutida en la Sección 3.5, está muy difundida en el campo de la voz humana y es conocida como separación fuente-filtro. Sin embargo, los tejidos vibrantes en el órgano vocal pueden presentar comportamiento irregular aún sin tener en cuenta una estructura intrínseca. Esto puede ocurrir, por ejemplo, si se considera la realimentación acústica retrasada, como se muestra a continuación.

5.1.4 Realimentación retrasada

El acople acústico entre la fuente sonora y el tracto vocal surge naturalmente cuando se tiene en cuenta que la onda de sonido creada por la fuente puede afectar a la misma fuente después de rebotar en su viaje a lo largo del tracto. En otras palabras, la fuente no “siente” una presión constante en el tracto, sino una que varía en el tiempo de acuerdo a los rebotes. El acople acústico entre fuente y tracto es una forma de *realimentación retrasada* (o *delayed feedback*), un mecanismo que le otorga soluciones complejas a sistemas dinámicos de baja dimensión [Ikeda and Matsumoto 1987, Hegger et al. 1998, Hale 1977].

Recientemente, la validez de la hipótesis de separación fuente-filtro fue demostrada experimentalmente en vocalizaciones tonales de palomas [Beckers et al. 2003a]. Sin embargo, no hay todavía pruebas experimentales de la validez de la hipótesis en vocalizaciones complejas como la mostrada en la Figura 5.1b, y no hay motivos para asumirla *a priori* en dichos casos. En efecto, la interacción entre fuente y tracto a través de realimentación acústica retrasada es un posible origen de dinámica no trivial en el órgano vocal, como mostramos en [Laje et al. 2001]. Allí estudiamos la oscilación de las cuerdas vocales humanas acopladas al tracto vocal con el modelo bidimensional de Titze (descrito en la Sección 3.3.3 y modificado para amplitud finita, ecuación 3.8). Mostramos que la interacción acústica retrasada con la onda de presión que se propaga en el tracto vocal le otorga a este sencillo sistema una variedad de soluciones subarmónicas, como se ve en la Figura 5.2.

En la siguiente sección presentamos un modelo para la interacción acústica entre fuente y tracto que nos permite estudiar fenómenos que no pueden ser descritos bajo la hipótesis de separación fuente-filtro, y realizamos predicciones refutables sobre el origen de la complejidad en el canto de algunas aves.

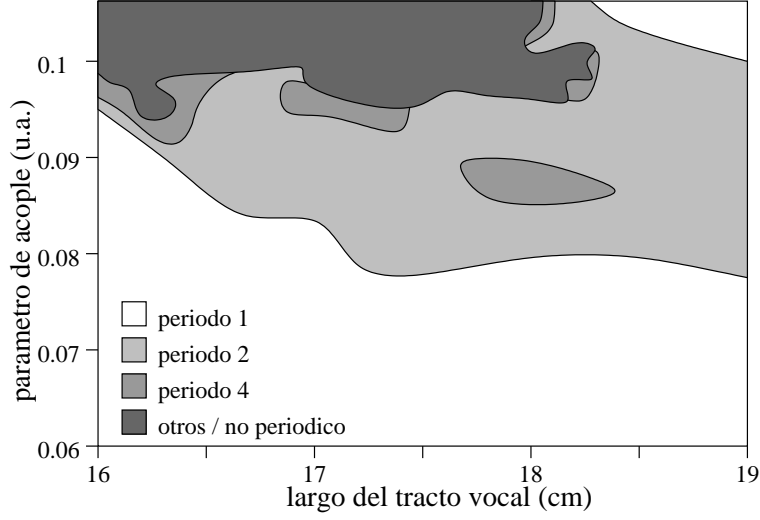


Figura 5.2: Diagrama de bifurcaciones esquemático de un modelo simple de las cuerdas vocales humanas, con realimentación acústica retrasada. Nótese que un incremento en el valor del parámetro de acople entre fuente y tracto provoca duplicaciones de período, un fenómeno típicamente no lineal.

5.2 Interacción acústica fuente-tracto

En esta sección presentamos un modelo para la producción vocal en Oscinos que permite estudiar la interacción acústica fuente-tracto [Laje and Mindlin 2005]. Aunque la siringe de los Oscinos tiene dos fuentes de sonido, muchas especies no utilizan los dos lados al mismo tiempo, y hay especies que utilizan prácticamente un solo lado. Es pertinente estudiar, por este motivo, el caso de una sola fuente acoplada al tracto.

Hemos dicho ya que un modelo del órgano vocal que considere la estructura intrínseca de los labia presenta por sí mismo fenómenos de dinámica no trivial, como el modelo de dos masas (ver Secciones 3.1.3 y 5.1.3). Para estudiar el efecto que la realimentación acústica pueda inducir en la dinámica de los labios debemos, por lo tanto, separar las fuentes de complejidad. Es necesario utilizar para la fuente sonora aislada un modelo que involucre un número reducido de modos de oscilación, y que por consiguiente tenga una dinámica simple (2D). Hechas estas consideraciones, utilizaremos para la oscilación labial el modelo bidimensional presentado en la Sección 3.3.4 (ecuaciones 3.11–3.13):

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x - by - cx^2y - f_0 + p_g,\end{aligned}\quad (5.3)$$

$$p_g = p_s + (Dy - 1)(p_s - p_i) \quad (5.4)$$

$$p_i = p_i(x, y). \quad (5.5)$$

En lo que sigue desarrollamos una expresión para $p_i = p_i(x, y)$ (es decir, una expresión para el acople acústico), a fin de obtener un modelo completo para el sistema acoplado fuente-tracto.

5.2.1 Acople acústico entre fuente y tracto

En el estudio de las cuerdas vocales humanas, el acople entre la fuente sonora y el tracto vocal fue históricamente considerado como dependiente del flujo $U(t)$ a través de la glotis y su derivada dU/dt , por medio de ciertos coeficientes de acople [Titze 1988, Titze and Story 1997, Fletcher and Tarnopolsky 1999]. Dichos coeficientes fueron tradicionalmente asociados a las partes real e imaginaria de la impedancia del tracto vocal, que típicamente depende de parámetros no locales como el largo y la sección del tracto vocal.

Este tradicional acercamiento se basa en la hipótesis de separación fuente-filtro: la dinámica de la fuente activa no lineal (la válvula vocal) es independiente de la onda de presión en el filtro pasivo lineal (el tracto vocal). La respuesta del filtro se representa por medio de la impedancia, que es el cociente entre la presión acústica y la velocidad de partícula (ambos complejos). Este cociente está bien definido sólo para sistemas lineales. Sin embargo, cuando fuente y filtro están acoplados e interactúan, el sistema completo es no lineal y el enfoque basado en la impedancia deja de ser válido [Mindlin and Laje 2005], como se describe en la Sección 3.5.

En un interesante trabajo sobre fonación humana [Story and Titze 1995], los autores consideran la propagación de la onda de presión en el tracto vocal, evitando así el uso de impedancias para modelar la respuesta del tracto. Sin embargo, restringieron el acople a una dependencia sólo de $U(t)$ y no de su derivada temporal.

Para tener en cuenta el efecto de la realimentación acústica en la dinámica del modelo, ecuaciones (5.3) y (5.4), en esta sección obtendremos una expresión para $p_i = p_i(x, y)$, esto es el acople fuente-tracto representado por la ecuación (5.5).

Dinámica de la onda acústica en el tracto

El panorama general de la dinámica de la onda de presión en el tracto vocal es como se describe a continuación. Una perturbación acústica es inyectada por la fuente y viaja como una onda de presión a lo largo del tracto. La onda sufre reflexiones parciales durante el viaje, en las interfases entre diferentes secciones del tracto y finalmente en el extremo opuesto. Las reflexiones se propagan hacia atrás hasta la base de la tráquea, afectando eventualmente un tiempo después la dinámica de la fuente. Esta dinámica puede ser tenida en cuenta por medio de condiciones de contorno muy simples (ver Sección 3.5) bajo la hipótesis de que la onda en el tracto es una onda plana, como sigue.

La presión a la entrada del tracto vocal p_i se compone de dos partes: las perturbaciones de la presión $s(t)$ generadas por el flujo $U(t)$ inyectadas localmente por la fuente, y una onda acústica $b(t)$ que se propaga hacia atrás debida a las reflexiones en el tracto:

$$p_i(t) = s(t) + b(t - T/2) \quad (5.6)$$

$$b(t) = -\gamma p_i(t - T/2), \quad (5.7)$$

donde $T/2 = L/c$ es el tiempo que le lleva a la onda sonora recorrer el tracto de largo L a una velocidad c , y γ es el coeficiente de reflexión en el extremo final del tracto (que por simplicidad consideramos como un tubo uniforme). La perturbación de la presión $s(t)$ inyectada por la fuente es una función de todas

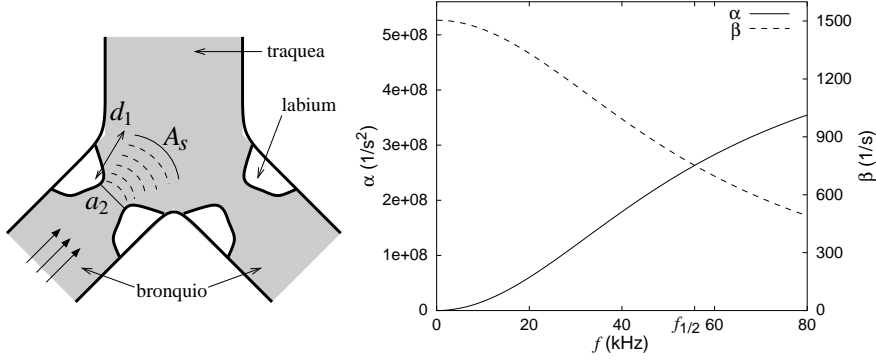


Figura 5.3: Acople acústico entre fuente y tracto. (*Izquierda*) Sección ventral esquemática de la siringe (no a escala). La modulación del flujo de aire por parte de los labia crea una onda acústica divergente que es inyectada en la base de la tráquea. Las líneas discontinuas indican la región de propagación esférica. (*Derecha*) Parámetros de acople α y β , como función de la frecuencia f (ecuaciones 5.14 y 5.15). A bajas frecuencias, el único parámetro que sobrevive es β . En este caso, se dice que el acople es *inertivo*, o sea que presión acústica y velocidad de partícula están prácticamente $\pi/2$ fuera de fase. En cambio, a altas frecuencias el parámetro más importante es α y el acople es *resistivo* (presión en fase con velocidad). Los valores de parámetros utilizados en este capítulo son los de la Tabla A.1, a menos que se indique lo contrario.

las variables que describen la dinámica de los labia. Una sustitución nos permite escribir las condiciones de contorno (5.6) y (5.7) más sucintamente:

$$p_i(t) = s(t) - \gamma p_i(t - T), \quad (5.8)$$

con T el tiempo de viaje ida y vuelta a lo largo del tracto. Las simples condiciones de contorno expresadas por la ecuación (5.8) asumen que la onda que se propaga en el tracto es una onda plana. Esta hipótesis es plausible en muchos tractos vocales de aves, puesto que sabemos por los fundamentos de acústica [Kinsler et al. 1982] que una guía de ondas de paredes rígidas cuyo diámetro sea mucho menor que la longitud de onda del sonido sólo propaga ondas planas. Una tráquea típica en Oscinos es un tubo de aproximadamente 1 mm de diámetro, mientras que la longitud de onda correspondiente a una frecuencia típica de 1000 Hz es de 35 cm.

Para continuar, es necesaria una expresión para $s(t)$ en función de las variables del labium, x e y . Es importante recalcar que la presión p_i está formada por la contribución *local* $s(t)$ (inyectada por la glotis en la base del tracto) y la contribución *retrasada* $p_i(t - T)$ en la forma de una onda plana.

Perturbaciones creadas por la fuente

Muy cerca de la fuente, sin embargo, no es de esperar que la onda sonora sea plana. Allí, la onda emitida es de naturaleza más bien divergente. Justamente, para obtener una expresión para la perturbación de la presión $s(t)$ inyectada por la válvula vocal, asumimos por simplicidad que la válvula es una fuente emisora de ondas armónicas localmente divergentes y esféricas (ver Figura 5.3a). La expresión buscada es una relación entre la presión $s(t)$ a la entrada del tracto

(al final de la zona de propagación esférica) y la velocidad de partícula $v(t)$ justo a la salida de los labia (que depende de las variables x e y). En la aproximación que la región de propagación esférica sea suficientemente pequeña, la relación buscada entre presión $s(t)$ y velocidad de partícula $v(t)$ para la onda esférica está dada por

$$s(t) = Rv(t) + I\dot{v}(t), \quad (5.9)$$

donde R e I son las resistencia e inercia acústicas específicas, respectivamente [Kinsler et al. 1982] (redefinidas en este trabajo por unidad de M). R e I son las partes real e imaginaria de la impedancia acústica específica $Z = R + i\omega I$. Los detalles de esta derivación pueden encontrarse en el Apéndice, Sección A.1. R e I están dadas por

$$R = \frac{\rho c}{M} \frac{(kr)^2}{1 + (kr)^2} \quad (5.10)$$

$$I = \frac{\rho}{M} \frac{r}{1 + (kr)^2}, \quad (5.11)$$

donde ρ es la densidad del aire y c la velocidad del sonido [Kinsler et al. 1982]. R e I dependen de la frecuencia del sonido f a través del número de onda $k = 2\pi f/c$.

Ahora necesitamos una expresión para $v(t)$ en términos de las variables x e y , a fin de relacionar la presión $s(t)$ con la dinámica de los labia. La conservación del flujo en la pequeña región de propagación esférica (Figura 5.3a) nos permite escribir

$$\begin{aligned} v(t) &= \frac{v_0}{A_s} a_2(t) \\ &= \frac{2hv_0}{A_s} (x_0 + x - \tau y), \end{aligned} \quad (5.12)$$

donde $a_2(t)$ es el área de salida de la válvula vocal (ecuación 3.3) y A_s el área del frente esférico a la entrada del tracto (ver detalles en el Apéndice, Sección A.1; v_0 es la velocidad media de partícula dada por Bernouilli $v_0 = \sqrt{2p_s M/\rho}$). Por último reemplazamos la ecuación (5.12) en la ecuación (5.9), redefinimos los coeficientes y tiramos el término constante (que no aporta a la presión acústica), para obtener

$$s(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}), \quad (5.13)$$

que nos permite completar una expresión para el acople entre fuente y tracto. Hemos definido los coeficientes de acople

$$\alpha = 2h \frac{v_0}{A_s} R \quad (5.14)$$

$$\beta = 2h \frac{v_0}{A_s} I, \quad (5.15)$$

que son simplemente un rescaleo de R e I , respectivamente, y que dependen de la frecuencia del sonido f (a través de R e I) en la forma mostrada en la Figura 5.3b. Los coeficientes α y β no son independientes, sino que están relacionados a través de

$$\alpha = F\beta, \quad (5.16)$$

con $F = 4\pi^2 f^2 d_1/c$.

En resumen, las ecuaciones (5.8) y (5.13) constituyen una expresión para la presión p_i a la entrada del tracto en función de las variables de la válvula vocal, es decir una expresión para el acople fuente-tracto:

$$p_i(t) = s(t) - \gamma p_i(t - T) \quad (5.17)$$

$$s(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}). \quad (5.18)$$

5.2.2 Modelo para una fuente acoplada al tracto

Hasta aquí, el modelo para una fuente acoplada al tracto vocal se forma con las ecuaciones (5.3), (5.4), (5.17) y (5.18):

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\epsilon x - by - cx^2y - f_0 + p_g \end{aligned} \quad (5.19)$$

$$p_g = p_s + (Dy - 1)(p_s - p_i) \quad (5.20)$$

$$p_i(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}) - \gamma p_i(t - T). \quad (5.21)$$

Nótese que la ecuación (5.21) introduce un término de aceleración \dot{y} en el miembro derecho de la ley de Newton, ecuación (5.19); dicho término es heredado directamente del modelo flameante y refleja el hecho de que el flujo que sale de la válvula no está en fase con x . Este término hace que el sistema pierda su forma explícita $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, lo que impide utilizar un integrador numérico estándar.

El sistema puede ponerse en forma explícita reordenando y reagrupando los términos con \dot{y} . No obstante, vale la pena considerar la aproximación de bajo acople $\beta \ll 1/\tau$, que no solamente deja al sistema en la forma explícita sino que además le da una forma mucho más simple. Los detalles de este cálculo se encuentran en la Sección A.2 del Apéndice. Considerando dicha aproximación, la forma final del modelo para una fuente acoplada al tracto es

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\epsilon x - by - cx^2y - f_0 + \tilde{p}_g \end{aligned} \quad (5.22)$$

$$\tilde{p}_g = p_s + (Dy - 1)(p_s - \tilde{p}_i) \quad (5.23)$$

$$\tilde{p}_i(t) = \alpha(x - \tau y) + \beta y - \gamma p_i(t - T) \quad (5.24)$$

$$p_i(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}) - \gamma p_i(t - T) \quad (5.25)$$

Nótese la diferencia entre las definiciones de p_i y \tilde{p}_i . Este detalle técnico (discutido en el Apéndice) permite utilizar una rutina estándar para integrar numéricamente el sistema.

Simulaciones

El modelo de una fuente acoplada al tracto con realimentación acústica, ecuaciones (5.22)–(5.25), presenta una dinámica muy rica. Observaciones preliminares muestran soluciones oscilatorias periódicas y no periódicas, y bifurcaciones de duplicación de período.

Como una ilustración de la variedad de soluciones, en la Figura 5.4 mostramos simulaciones para dos posibles maniobras vocales: (a) variación de ϵ , que

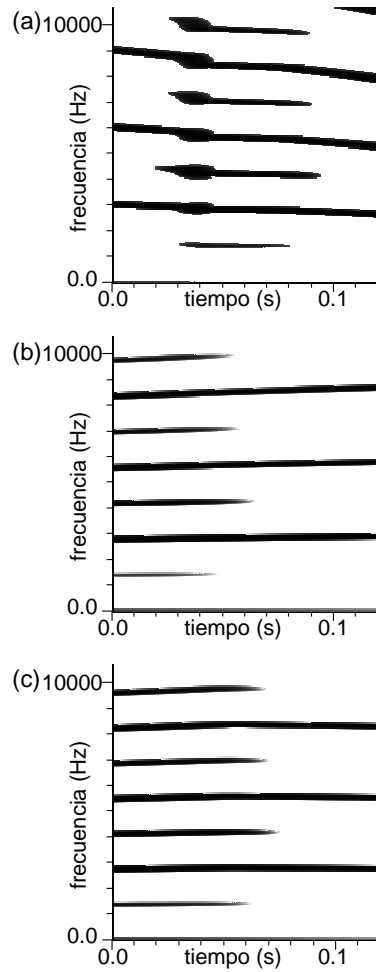


Figura 5.4: Simulaciones del modelo de una fuente acoplada al tracto con realimentación acústica retrasada, ecuaciones (5.22)–(5.25): sonogramas de la presión acústica p_i para diferentes maniobras vocales. Los valores de parámetros son los de la Tabla A.1, a menos que se indique lo contrario. (a) Variando ϵ . Frecuencias subarmónicas aparecen y desaparecen cuando ϵ decrece continuamente desde $3.6 \times 10^8 \text{ s}^{-2}$ a $2.6 \times 10^8 \text{ s}^{-2}$. (b) Variando p_s . Frecuencias subarmónicas desaparecen cuando p_s decrece continuamente desde $6.2 \times 10^6 \text{ cm s}^{-2}$ a $4.2 \times 10^6 \text{ cm s}^{-2}$. (c) Variando β (y α , de acuerdo con la ecuación 5.16). Las frecuencias subarmónicas desaparecen cuando β disminuye continuamente desde 1900 s^{-1} a 900 s^{-1} .

se interpreta como una variación en la tensión del músculo vS y provoca una variación en la frecuencia fundamental de la vocalización; y (b) variación de p_s , que se interpreta como una variación en la presión del saco aéreo y provoca una variación en la intensidad del sonido. La interpretación de ambas maniobras en términos de los parámetros fisiológicos fue establecida en el Capítulo 4. Además de las dos maniobras vocales referidas, se muestra también una tercera simulación variando β (y α , de acuerdo con la ecuación 5.16), para estudiar el efecto de diferentes configuraciones de la interfase fuente-tracto en la dinámica del sistema. Recordemos que β depende de parámetros anatómicos como d_1 , que puede interpretarse como una medida de la distancia entre la válvula vocal y la base de la tráquea (ver ecuación 5.15 y Figura 5.3).

Duplicaciones de período y de frecuencia son evidentes en la Figura 5.4. Vale la pena recalcar que la realimentación retrasada es el único origen dinámico de este rico comportamiento, dado que la fuente está modelada como un sistema dinámico bidimensional; no puede tener comportamiento complejo por sí sola. En efecto, para valores bajos de los parámetros de acople se ve que el compor-

tamiento del sistema es bidimensional.

5.2.3 Predicciones

Una manera de explorar y validar este modelo consiste en estudiar cómo la dinámica es afectada cuando ciertos parámetros son variados. En particular, la densidad de la atmósfera en la cual opera el modelo puede ser modificada. Nuestro modelo predice que, a la vista de un sonograma subarmónico, observaremos cambios cualitativos (o no) cuando la densidad de la atmósfera cambie, dependiendo de cuál sea el origen de la subarmonicidad: interacción acústica entre fuente y tracto, o modos altos de vibración de los labia. En lo que sigue mostraremos que los parámetros de acople α y β cambian apreciablemente cuando el ave canta en una atmósfera de heliox, en la cual el nitrógeno, que compone el 80% del aire normal, es reemplazado por el menos denso helio. La densidad del helio es $1/8$ la del nitrógeno, por lo cual el heliox tiene una densidad 30% la del aire normal. La velocidad del sonido en heliox es 74% mayor que en aire [Gray 1963, Nowicki 1987].

Experimentos con heliox fueron utilizados para mostrar que algunas aves coordinan activamente las características del filtrado pasivo de su tracto vocal con la salida de la siringe [Nowicki 1987]. En ese trabajo se mostró que el tracto modula las características espectrales del sonido producido en la siringe, pero que tiene poca influencia en la vibración de las membranas. El heliox fue también utilizado para probar la hipótesis del silbido en loros [Brittan-Powell et al. 1997] y palomas [Ballintijn and ten Cate 1998], como se describe en la Sección 3.2.1.

El resultado esperado de cada uno de los experimentos descritos en el párrafo anterior fue tanto un cambio en la frecuencia fundamental de la vocalización (en la misma proporción en que cambia la velocidad del sonido), o un cambio sólo en las resonancias del tracto con poca variación de la frecuencia fundamental. Sin embargo, en esta sección mostraremos que otros posibles resultados pueden esperarse de un experimento con heliox cuando la fuente está acústicamente acoplada con el tracto. Dichos resultados son diferentes de aquellos esperados en los dos escenarios de acople tradicionalmente considerados, donde la frecuencia fundamental cambia en la misma proporción que las resonancias del tubo: un silbido [Ballintijn and ten Cate 1998, Nowicki 1987, Brittan-Powell et al. 1997], o una válvula vibrante controlada por presión [Fletcher and Tarnopolsky 1999, Brittan-Powell et al. 1997] como la de los instrumentos de lengüeta o los bronces.

Escalear de los coeficientes de acople

En lo que sigue haremos explícita la dependencia de los coeficientes de acople α y β respecto de los parámetros acústicos ρ (densidad del aire) y c (velocidad del sonido). Recordemos primero la definición de la inercia acústica específica I , ecuación (5.11). Podemos escribir

$$\begin{aligned} I &= \frac{\rho}{M} \frac{d_1}{1 + (kd_1)^2} \\ &\sim \rho g(f), \end{aligned} \tag{5.26}$$

con

$$g(f) = \frac{1}{1 + (kd_1)^2}, \quad k = 2\pi f/c.$$

Notemos que la función $g(f)$ depende de c a través del número de onda k . Sin embargo, $g(f) \sim 1$ a bajas frecuencias y cambia muy poco con c cuando el aire normal es reemplazado por heliox. Con esto y la ecuación (5.15), podemos escribir explícitamente la dependencia de β con los parámetros acústicos:

$$\begin{aligned}\beta &= 2h \frac{v_0}{A_s} I \\ &\sim \rho^{1/2} g(f),\end{aligned}\tag{5.27}$$

donde usamos que $v_0 \sim \rho^{-1/2}$ (ver la definición que sigue a la ecuación 5.12). Para obtener la dependencia de α con los parámetros acústicos, recordemos que α y β están relacionados por la ecuación (5.16). Por lo tanto,

$$\begin{aligned}\alpha &= 4\pi^2 f^2 d_1 c^{-1} \beta \\ &\sim \rho^{1/2} c^{-1} g(f).\end{aligned}\tag{5.28}$$

Hasta acá, las ecuaciones (5.27) y (5.28) nos dan la dependencia explícita respecto de los parámetros acústicos ρ y c . Teniendo en cuenta el comportamiento de $g(f)$ discutido más arriba, podemos decir que al cambiar de aire normal (a) a heliox (h) los coeficientes de acople deberían escalear aproximadamente como

$$\begin{aligned}\alpha_h &\sim \frac{c_a}{c_h} \left(\frac{\rho_a}{\rho_h} \right)^{-1/2} \alpha_a \\ &\sim 0.31 \alpha_a\end{aligned}\tag{5.29}$$

$$\begin{aligned}\beta_h &\sim \left(\frac{\rho_a}{\rho_h} \right)^{-1/2} \beta_a \\ &\sim 0.55 \beta_a.\end{aligned}\tag{5.30}$$

Predicciones

Este resultado sugiere que pueden ocurrir cambios perceptibles en los sonogramas cuando el aire es reemplazado por heliox. En la Figura 5.5a mostramos una simulación del modelo, ecuaciones (5.22)–(5.25). El parámetro p_s decrece exactamente como en la Figura 5.4b. En el recuadro de la izquierda (aire normal), el comportamiento subarmónico es evidente al comienzo del sonograma. Las frecuencias subarmónicas desaparecen en una bifurcación de duplicación de período inversa alrededor de 0.05 s cuando p_s cruza por debajo del valor de bifurcación. Cuando el aire es reemplazado por heliox (recuadro de la derecha), las frecuencias subarmónicas no aparecen.

Vale la pena notar que, contrariamente a lo que podría esperarse en otros mecanismos de válvula vibrante acoplada a un resonador, como una trompeta, un trombón o instrumentos de lengüeta [Brittan-Powell et al. 1997], la frecuencia fundamental casi no cambia al pasar de aire a heliox. Sin embargo el acople está presente, puesto que son evidentes los cambios en el sonograma.

Por el contrario, un experimento de heliox mostrará pocas diferencias si el origen dinámico de las frecuencias subarmónicas en el sonograma es la aparición de modos altos de vibración de los labia. En la Figura 5.5b mostramos una simulación de un modelo de dos masas asimétrico sin acople a un tracto, que no depende de la densidad del aire o de la velocidad del sonido [Steinecke and

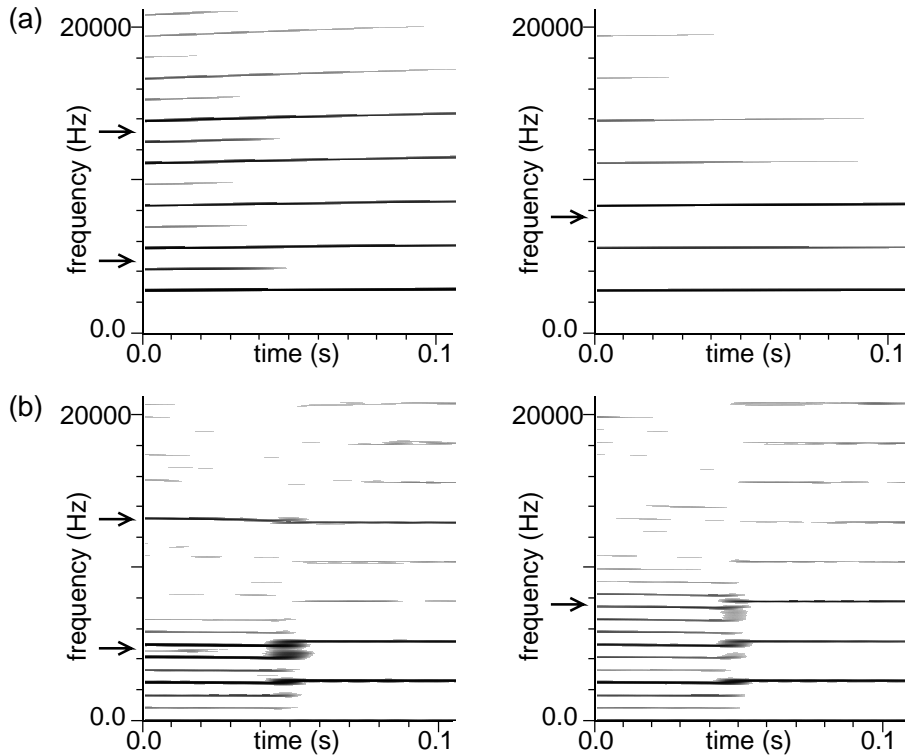


Figura 5.5: Sonogramas esperados en experimentos de heliox. Los valores de parámetros son los de la Tabla A.1, a menos que se indique lo contrario. (a) Simulación del modelo de una fuente acoplada al tracto con realimentación acústica retrasada, ecuaciones (5.22)–(5.25). El parámetro p_s decrece exactamente como en la Figura 5.4b (aunque hay diferencias en los tonos de grises, debidas a la escala lineal de intensidad utilizada aquí para permitir la discriminación de las resonancias del tracto). (*Izquierda*): Aire normal, al comienzo del sonograma se observan frecuencias subarmónicas que desaparecen en una bifurcación alrededor de 0.05 s. (*Derecha*): Misma maniobra, pero en heliox. El comportamiento subarmónico no está presente. Sin considerar la bifurcación, vale la pena notar que, aunque fuente y tracto están acoplados, la frecuencia fundamental no varía al cambiar de aire a heliox. (b) Simulación del modelo de dos masas asimétrico [Steinecke and Herzel 1995], sin acople a un tracto (la serie temporal de la presión es filtrada *a posteriori*, según la ecuación 5.17). Los valores de las masas y los coeficientes de restitución elástica del lado izquierdo del modelo son el 52% de los del lado derecho. La presión sublabial decrece en forma análoga a la de (a). (*Izquierda*): Aire normal. Alrededor de 0.05 s ocurre una triplicación de período inversa. (*Derecha*): Misma maniobra, pero en heliox. El comportamiento subarmónico no desaparece. El único cambio perceptible es un incremento en las frecuencias de resonancia del tracto vocal. Las flechas indican aproximadamente las resonancias del tracto: 4375 Hz y 13125 Hz (aire normal), y 7612 Hz (heliox). Resonancias más altas caen más allá de 20000 Hz.

Herzel 1995]. Se trata de una maniobra vocal similar a la de la Figura 5.5a: la presión sublabial decrece, provocando en este caso una triplicación de período inversa. Las frecuencias subarmónicas están presentes, sin embargo, tanto en aire como en heliox. Queremos destacar que las únicas diferencias entre recuadros izquierdos y derechos correspondientes son los valores de densidad del aire y velocidad de sonido.

No obstante, podremos afirmar algo sobre el efecto de la realimentación acústica sólo en el caso en que encontremos una diferencia como la mostrada en la Figura 5.5a. Para un ave en particular, puede ocurrir que la variación en la densidad del aire y la velocidad del sonido lograda al cambiar a heliox no sea suficiente para disminuir los coeficientes de acople al punto que el efecto de la realimentación acústica sea suprimido. Otras maneras de modificar la realimentación acústica deberían ser ensayados en tal caso, por ejemplo modificando el largo efectivo del tracto vocal por medio de una cánula insertada en alguna posición en la tráquea (efectos posibles de una traqueotomía se discuten en [Beckers et al. 2003a]), u otras modificaciones mecánicas o anatómicas.

5.3 Interacción acústica fuente-fuente

El aparato vocal de las aves despliega una increíble capacidad para producir sonidos. A la riqueza de vocalizaciones producidas por un solo lado de la siringe (como las estudiadas en la sección anterior y en el Capítulo 3) debemos agregar las producidas por el otro lado, dado que muchas especies tienen la capacidad de cantar con los dos lados al mismo tiempo en forma independiente, como en el sonograma de la Figura 5.1a. La estructura bilateral de la siringe, sin embargo, nos depara mucho más que el doble de lo que ya vimos.

5.3.1 Fuentes interactuantes

Huellas espectrales de interacción

La huella espectral de la interacción entre dos sistemas periódicos es considerada tradicionalmente en términos de *modulación* [Elemans 2004]. Esto es, si la evolución individual de uno de los sistemas es descrita por la señal $f(t) = A \cos(\omega t + \phi)$, la interacción con el otro sistema se refleja en la modulación periódica de uno o varios de los parámetros de $f(t)$:

$$f(t) = A(t) \cos[\omega(t)t + \phi(t)],$$

llamadas modulación de amplitud, de frecuencia y de fase, respectivamente. Consideremos por ejemplo el caso simple de modulación de la amplitud de una función armónica. La siguiente función

$$f(t) = A \cos \omega_1 t \tag{5.31}$$

tiene un sonograma que se compone, como era de esperarse, por un único trazo de frecuencia $F_1 = \omega_1/2\pi$. Ahora consideremos que el sistema que produce la señal $f(t)$ interactúa de alguna forma con un agente externo, de manera que se refleja en una modulación de la amplitud a una frecuencia F_2 : $A = A(t) =$

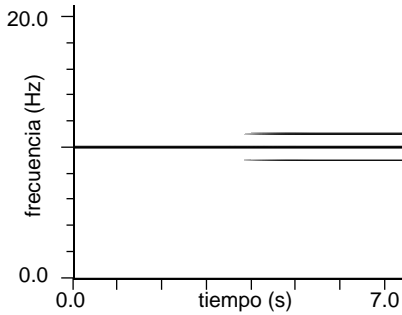


Figura 5.6: Modulación de amplitud de una señal armónica. El sonograma corresponde a una señal con una frecuencia de 10 Hz. La amplitud de la señal es modulada súbitamente a partir de los 4 s con una frecuencia de 1 Hz, lo que resulta en la aparición de trazos paralelos al original en 9 y 11 Hz.

$A_0 + \epsilon \cos \omega_2 t$. En este caso, podemos reescribir la ecuación anterior como

$$f(t) = A(t) \cos \omega_1 t \quad (5.32)$$

$$= (A_0 + \epsilon \cos \omega_2 t) \cos \omega_1 t$$

$$= A_0 \cos \omega_1 t + \frac{1}{2} \epsilon \cos(\omega_1 - \omega_2)t + \frac{1}{2} \epsilon \cos(\omega_1 + \omega_2)t, \quad (5.33)$$

donde utilizamos la identidad trigonométrica $\cos \alpha \cos \beta = \frac{1}{2}[\cos(\alpha - \beta) + \cos(\alpha + \beta)]$. Notemos que la modulación de la amplitud provoca la aparición de “nuevas frecuencias” $\omega_1 + \omega_2$ y $\omega_1 - \omega_2$, que en el sonograma aparecen como bandas paralelas al trazo original (ver Figura 5.6). En general, las nuevas frecuencias no están relacionadas armónicamente con la frecuencia original.

Evidencia experimental de interacción fuente-fuente en aves

La interacción entre las dos fuentes de sonido de un ave fue demostrada experimentalmente por primera vez por Nowicki y Capranica en el paro de corona negra (*Parus atricapillus*) [Nowicki and Capranica 1986a, Nowicki and Capranica 1986b]. Incapacitando diferentes lados de la siringe en varias aves, los autores demostraron que el espectro complejo de una de las sílabas era más que la simple superposición de los espectros de las fuentes por separado: aparecían frecuencias sumas y diferencias (Figura 5.7). Las causas físicas de la interacción, sin embargo, fueron especulativas, aunque concluyeron que la interacción era pasiva.

Un hecho asombroso es que algunas especies tienen la capacidad de elegir entre usar sus dos fuentes en forma independiente o hacerlas interactuar. Este parece ser el caso del tangará overo o moriche overo (*Cissopis leveriana*) [Straneck 1990c]. El sonograma de una de sus sílabas puede verse en la Figura 5.8. Los trazos a la izquierda de la línea punteada ponen en evidencia que las dos fuentes están actuando en forma independiente: no están relacionados temporal ni armónicamente. A la derecha de la línea punteada, sin embargo, la cosa cambia: aunque los trazos no tienen relación armónica entre ellos, la evolución temporal es la misma hasta en los más mínimos detalles, y además aparecen trazos paralelos a los dos originales. El ave ha modificado su órgano vocal y/o su tracto para hacer que las fuentes interactúen. No podemos decir nada sobre el mecanismo de interacción, pero podemos estar seguros de que las fuentes están interactuando.

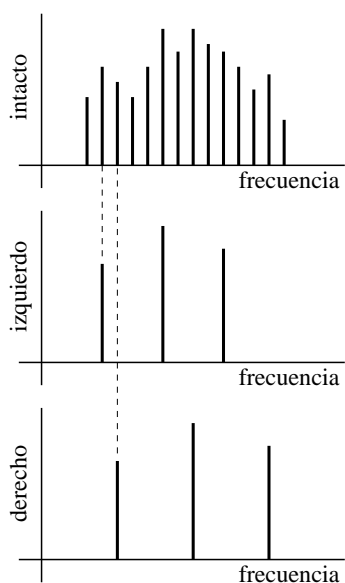


Figura 5.7: Interacción entre las dos fuentes en el paro de corona negra (*Parus atricapillus*). Representación esquemática de los resultados presentados en [Nowicki and Capranica 1986a, Nowicki and Capranica 1986b]. (*Arriba*): Espectro de una sílaba del ave intacta. Las frecuencias presentes, que podrían ser armónicos de una fundamental ausente, no son necesariamente un múltiplo de la diferencia entre ellas. (*Centro, Abajo*): Espectros de cada una de las fuentes por separado. El espectro de la siringe intacta es la superposición de estos dos espectros, con el agregado de frecuencias sumas y diferencias que no están presentes en los espectros de cada lado.

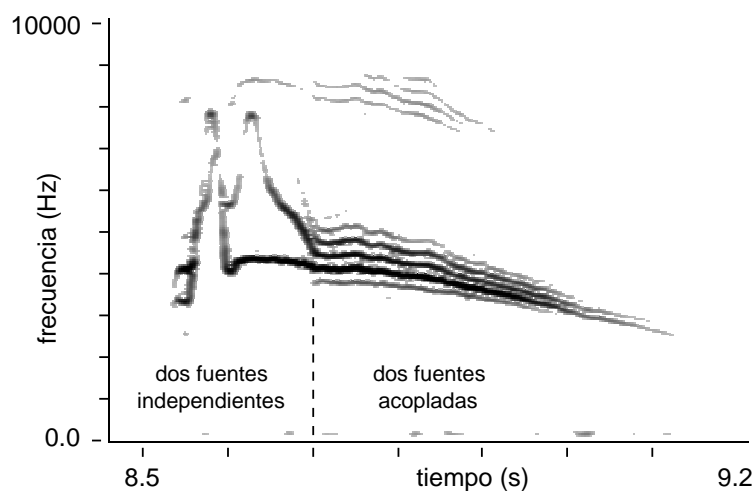


Figura 5.8: Una sílaba del canto del tangará overo (*Cissopis leveriana*). Al comienzo, la evolución de los trazos nos indica que las dos fuentes están actuando independientemente. Cuando las frecuencias fundamentales se acercan lo suficiente, alrededor de los 8.7 s, se puede observar la aparición de una serie de frecuencias subarmónicas. Es probable que algún tipo de acople entre las fuentes esté actuando aquí.

Interacción acústica

La interacción entre los lados de la siringe en el paro de corona negra puede tener diversos orígenes físicos. Los autores señalan que "... dado el pequeño tamaño del órgano vocal, la proximidad de los dos lados y las muchas conexiones físicas entre ellos, sería sorprendente que los dos lados nunca funcionaran de alguna forma acoplada..." [Nowicki and Capranica 1986b]. Una de las formas de interacción que arriesgan es el acople estructural, por ejemplo a través del tejido cartilaginoso que separa las paredes mediales de los bronquios. La otra forma de interacción considerada es una de tipo acústico, transmitida de una fuente a la otra a través del tracto vocal o del saco aéreo interclavicular.

La interacción acústica entre las fuentes surge naturalmente cuando se cae en la cuenta de que, en ciertas circunstancias, la presión en la base de la tráquea puede influir en la dinámica de una fuente individual, como se describe en la primera sección de este capítulo. Puede darse el caso, por lo tanto, que las fluctuaciones en la presión inducidas por una de las fuentes afecte a la otra. En lo que sigue presentamos un modelo para estudiar este tipo de comportamientos.

5.3.2 Modelo para dos fuentes acopladas al tracto

El modelo presentado en la Sección 5.2.2 para una sola fuente acoplada acústicamente al tracto vocal puede ampliarse para tener en cuenta las dos fuentes, identificadas a partir de ahora por los subíndices l (izquierda) y r (derecha).

Para empezar tenemos que tomar las ecuaciones (5.3) y (5.4), y escribir una ley de Newton para cada fuente:

$$\begin{aligned}\dot{x}_l &= y_l \\ \dot{y}_l &= -\epsilon_l x_l - b_l y_l - c_l x_l^2 y_l - f_{0l} + p_{gl}\end{aligned}\quad (5.34)$$

$$\begin{aligned}\dot{x}_r &= y_r \\ \dot{y}_r &= -\epsilon_r x_r - b_r y_r - c_r x_r^2 y_r - f_{0r} + p_{gr}\end{aligned}\quad (5.35)$$

$$p_{gl} = p_s + (D_l y_l - 1)(p_s - p_i) \quad (5.36)$$

$$p_{gr} = p_s + (D_r y_r - 1)(p_s - p_i). \quad (5.37)$$

Lo que sigue es notar que la presión p_i en la base del tracto vocal es común a ambas fuentes, puesto que en ese lugar confluyen los bronquios. Por lo tanto escribimos, indistintamente,

$$p_i(t) = s(t) - \gamma p_i(t - T). \quad (5.38)$$

En cambio, las perturbaciones de la presión $s(t)$ inyectadas por las fuentes en la base del tracto vocal deben tener en cuenta la contribución independiente de cada fuente. Esto es, la ecuación (5.18) debe modificarse para incorporar la contribución de cada fuente:

$$\begin{aligned}s(t) &= s_l(t) + s_r(t) \\ &= \alpha_l(x_l - \tau_l y_l) + \beta_l(y_l - \tau_l \dot{y}_l) \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r(y_r - \tau_r \dot{y}_r).\end{aligned}\quad (5.39)$$

Notemos que, gracias a esta expresión, las fuentes quedan acopladas a través de p_i .

Como fue notado en su momento en la Sección 5.2.2, la ecuación recién escrita introduce términos de aceleración \dot{y}_l, \dot{y}_r en los miembros derechos de las ecuaciones (5.34) y (5.35). Dichos términos hacen que el sistema pierda la forma explícita $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, y por lo tanto impiden el uso de un integrador numérico estándar.

Es posible, sin embargo, poner al sistema en forma explícita reagrupando términos. Es posible además simplificarlo considerablemente si adoptamos la aproximación $\beta_j \ll 1/\tau_j$ ($j = l, r$), análogamente a lo propuesto en la Sección 5.2.2. Los detalles de estos cálculos pueden consultarse en la Sección A.3 del Apéndice. Habiendo dicho esto, la forma final para nuestro modelo de dos fuentes acopladas acústicamente al tracto vocal con realimentación retrasada queda [Laje and Mindlin 2005]

$$\begin{aligned} \dot{x}_l &= y_l \\ \dot{y}_l &= -\epsilon_l x_l - b_l y_l - c_l x_l^2 y_l - f_{0l} + \tilde{p}_{gl} \end{aligned} \quad (5.40)$$

$$\begin{aligned} \dot{x}_r &= y_r \\ \dot{y}_r &= -\epsilon_r x_r - b_r y_r - c_r x_r^2 y_r - f_{0r} + \tilde{p}_{gr} \end{aligned} \quad (5.41)$$

$$\tilde{p}_{gl} = p_s + (D_l y_l - 1)(p_s - \tilde{p}_i) \quad (5.42)$$

$$\tilde{p}_{gr} = p_s + (D_r y_r - 1)(p_s - \tilde{p}_i) \quad (5.43)$$

$$\begin{aligned} \tilde{p}_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l y_l \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r y_r - \gamma p_i(t - T) \end{aligned} \quad (5.44)$$

$$\begin{aligned} p_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l(y_l - \tau_l \dot{y}_l) \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r(y_r - \tau_r \dot{y}_r) - \gamma p_i(t - T) \end{aligned} \quad (5.45)$$

La diferencia entre las definiciones de p_i y \tilde{p}_i permite resolver numéricamente el sistema con una herramienta de integración estándar del tipo Runge-Kutta (ver discusión en el Apéndice, Sección A.3).

Este modelo de dos fuentes acopladas se reduce al modelo de una sola fuente (desarrollado en la sección anterior, ecuaciones 5.22–5.25) al poner los coeficientes de acople de la segunda fuente iguales a cero. Poniendo, además, los coeficientes de acople de la primera fuente iguales a cero, se recupera el caso de una fuente aislada (ecuaciones 5.3–5.5, con $p_i = 0$). Como se discute al comienzo de este capítulo, el modelo de una fuente aislada es un sistema dinámico bidimensional, y como tal no tiene dinámica compleja por sí mismo. Por esta razón elegimos el modelo flameante para describir la vibración de los labia. De esta manera los posibles orígenes de la complejidad están reducidos a los efectos del acople, lo que permite estudiar el efecto de la realimentación retrasada en la dinámica de una fuente, como también el efecto de acoplar acústicamente las dos fuentes.

5.3.3 Simulaciones

En la Figura 5.9a se muestra una simulación del modelo presentado en la sección anterior (un estudio detallado de la dinámica de este modelo será publicado más

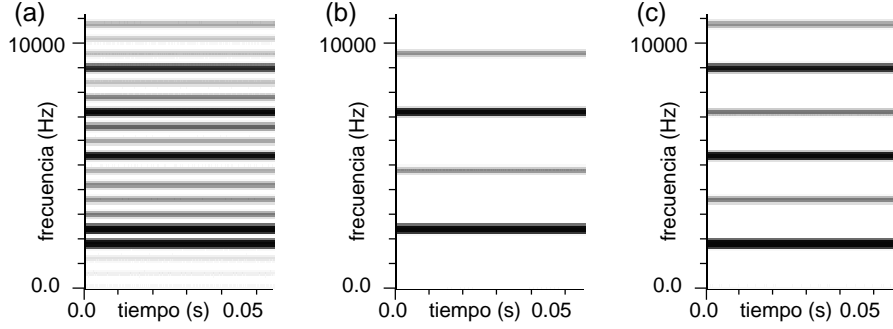


Figura 5.9: Simulación utilizando el modelo de dos fuentes acopladas al tracto, ecuaciones (5.40)–(5.45), sin tener en cuenta el efecto de la realimentación retrasada (la realimentación se suprime simplemente poniendo $\gamma = 0$). Los coeficientes elásticos de cada fuente en esta figura son $\epsilon_l = 2.4 \times 10^8 \text{ s}^{-2}$ y $\epsilon_r = 1.4 \times 10^8 \text{ s}^{-2}$. Todos los otros parámetros son los de la Tabla A.1. (a) Sonograma de la presión acústica p_i para las dos fuentes acopladas. Notar los dos trazos oscuros alrededor de 2000 Hz. Cada frecuencia en este sonograma es suma o diferencia de múltiplos de esas dos frecuencias. (b,c) Sonograma de las fuentes izquierda y derecha aisladas, respectivamente. Los valores de parámetros son iguales que en (a), salvo por los coeficientes de acople: en (b) ponemos $\alpha_r = \beta_r = 0$ para considerar sólo la fuente l (aislada de la fuente r). Recíprocamente, en (c) ponemos $\alpha_l = \beta_l = 0$ para considerar sólo la fuente l . Notar que el sonograma en (a) no es simplemente la superposición de los sonogramas de las fuentes aisladas, sino que tiene además varias otras frecuencias que son sumas y diferencias de múltiplos de las frecuencias fundamentales de las fuentes aisladas.

adelante [Sciamarella et al. 2005]). Para estudiar el efecto que el acople entre las dos fuentes otorga a las soluciones del modelo, desactivamos la realimentación retrasada poniendo $\gamma = 0$.

En la misma Figura, paneles (b,c), se muestra por separado el espectro de las fuentes aisladas. Estos sonogramas fueron obtenidos al aislar una fuente de la otra, poniendo $\alpha_l = \beta_l = 0$ en un caso y $\alpha_r = \beta_r = 0$ en el otro. Notemos que hay frecuencias que aparecen en el panel (a), cuando ambas fuentes están presentes, que no aparecen en los espectros de las fuentes aisladas. Estas frecuencias son las frecuencias heterodinas

$$f_{mn} = mF_l + nF_r \quad (m, n \in Z), \quad (5.46)$$

que son sumas y diferencias de múltiplos de las frecuencias fundamentales de cada fuente, F_l y F_r .

El efecto de la realimentación retrasada fue desactivado a propósito poniendo $\gamma = 0$. Por lo tanto, el acople acústico entre las dos fuentes es el único origen dinámico de las frecuencias subarmónicas que aparecen en la Figura 5.9a. El hecho de que las fuentes aisladas no muestran dinámica compleja por sí mismas está ilustrado en los paneles (b) y (c).

Capítulo 6

Ritmos complejos

Now I will do nothing but listen...

Walt Whitman, Song of Myself (1855)

Hay mucho más en el canto de las aves que simplemente sonidos bellos o raros. Desde Clément Janequin (ca. 1485–1558) hasta Olivier Messiaen (1908–1992), muchos compositores le prestaron atención a las melodías que las aves nos proponen en sus cantos. Nosotros, hasta ahora, hemos puesto el foco apenas en las propiedades acústicas de las sílabas, que son los elementos individuales que componen la melodía. En este capítulo analizaremos otro aspecto de la melodía de las aves: el *ritmo*. Usamos *ritmo* en un sentido muy musical: queremos decir la “proporción guardada entre el tiempo de un movimiento y el de otro diferente” [DRAE], o el tiempo de los sonidos dentro de la melodía. Analizamos el canto a dúo del hornero: macho y hembra alternan sílabas de una manera aparentemente descoordinada. Sin embargo, el llamativo ritmo que forman entre los dos nos permite intuir características del cerebro de la hembra, pues la forma en que se acoplan las voces es la de un oscilador no lineal forzado [Laje and Mindlin 2003].

6.1 Los horneros

El hornero (*Furnarius rufus*, suborden Suboscinos) es un ave sudamericana mucho menos conocida por su canto que por su asombroso nido. El nido, que es construido todos los otoños al final de la época de lluvias, tiene forma de horno y está hecho de barro y paja. El hornero es un ave no migratoria y forma pareja estable para toda la vida; estas características, sumadas a sus habilidades constructoras, le valieron el título no oficial de “Ave de la Patria”, otorgado en el año 1928 cuando el diario La Razón de Buenos Aires hizo una encuesta entre niños de escuelas primarias de todo el país.

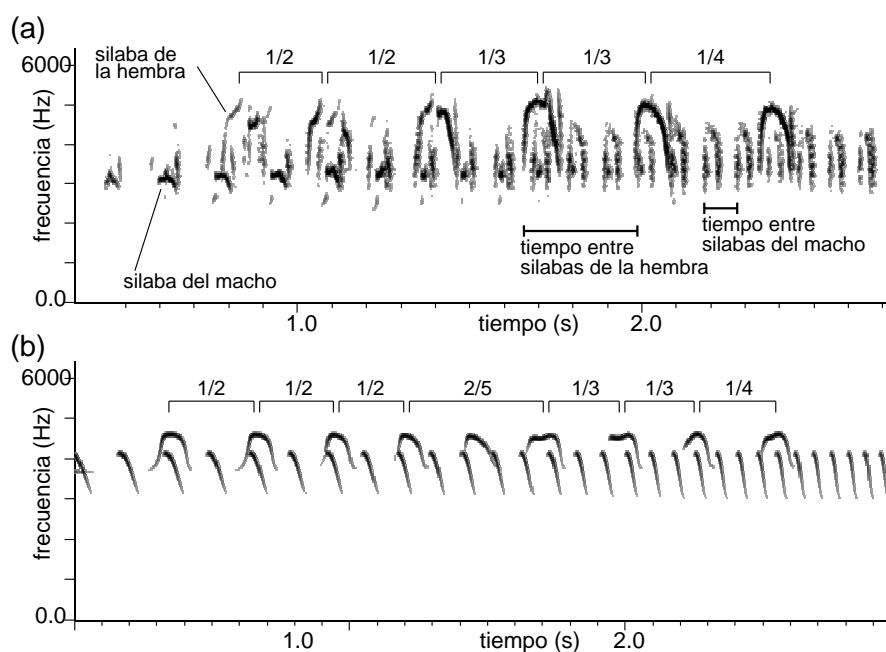


Figura 6.1: Sonogramas de dúos. (a) Sonograma de un típico dúo de horneros (*Furnarius rufus*). Las sílabas de macho y hembra son fácilmente distinguibles, debido a las diferentes características temporales y espectrales. Al comienzo, macho y hembra están sincronizados (una sílaba de la hembra por cada sílaba del macho), pero a poco de comenzar la hembra parece retrasarse. Sin embargo, hay una organización sistemática de las secuencias de acople que se muestran sobre el sonograma. (b) Sonograma de un dúo simulado. Las sílabas del macho fueron simuladas integrando el modelo de la siringe (ecuación 6.2) con funciones $\epsilon(t)$ y $p(t)$ armónicas de frecuencia creciente (ecuación 6.3). Las sílabas de la hembra también fueron simuladas integrando el modelo (ecuación 6.2), pero con funciones $\epsilon(t)$ y $p(t)$ dadas por un oscilador de relajación (ecuación 6.4) forzado a su vez por los gestos del macho.

6.1.1 El dúo de horneros

Lo que mucha gente no sabe es que, además de construir nidos de barro, los horneros cantan a dúo¹. El sonograma de un dúo típico (que puede durar hasta 10 s) puede verse en la Figura 6.1a. Las sílabas de macho y hembra son claramente distinguibles. El macho empieza cantando a una repetición de aproximadamente 5 Hz, y en algunos segundos incrementa la frecuencia de repetición hasta aproximadamente 20 Hz. Al comienzo del dúo la hembra parece seguir al macho, cantando una sílaba cada vez que el macho lo hace, en un modo “uno a uno”. Sin embargo, a poco de comenzar la hembra parece perder sincronización y retrasarse.

Mientras el macho en general aumenta la frecuencia de repetición a medida que el dúo se desarrolla, la hembra no tiene una tendencia definida: hay hembras que aumentan la repetición, otras disminuyen, y otras muestran comportamiento

¹Es llamativo que el canto a dúo, común en especies tropicales, se da habitualmente en especies que forman pareja estable [Burger 1980].

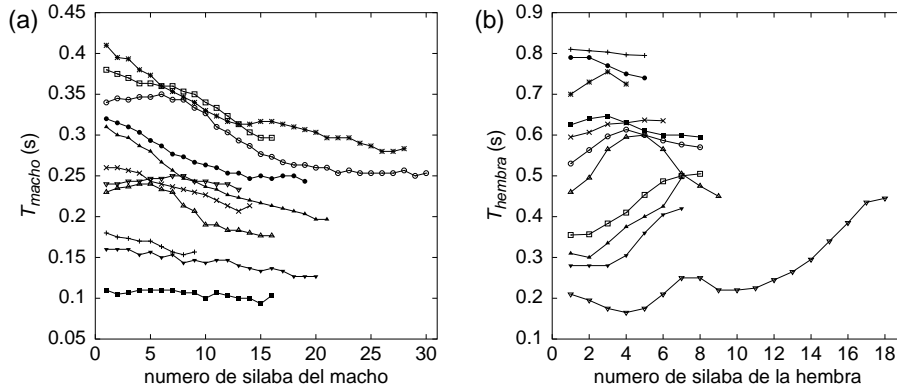


Figura 6.2: Comportamiento individual: tiempo promedio entre sílabas T como función del número de sílaba. (a) Los machos muestran una clara tendencia a aumentar la frecuencia de repetición a medida que el dúo se desarrolla. (b) Las hembras, en cambio, pueden aumentar, disminuir o variar no monóticamente la frecuencia de repetición. Las series temporales están desplazadas verticalmente por claridad.

mixto (Figuras 6.2a y b).

Tradicionalmente suelen distinguirse tres categorías de canto a dúo en aves [Burger 1980]: el canto *antifonal* (las frases o sílabas se emiten alternativamente), el canto *a dúo* (los miembros de la pareja emiten frases distintas simultáneamente) y el canto *a dúo simultáneo* (frases idénticas al unísono). El canto a dúo de los horneros no pertenece a ninguna de estas clases, pero está gobernado por una regla matemática subyacente como veremos a continuación.

6.1.2 Comportamiento conjunto

Para analizar el canto de los horneros, grabamos dúos de horneros en la Estación de Cría de Animales Silvestres (ECAS) en Villa Elisa, Argentina. A fin de estudiar el comportamiento conjunto de la pareja en un dúo, definimos una *coincidencia* como el evento en el cual el máximo de una sílaba del macho está abarcado por una sílaba de la hembra. Con esto, definimos un número r_{ap} que describe la cantidad de sílabas que ocurren entre coincidencias consecutivas. Este número es el cociente $r_{ap} = p/q$ entre los enteros p y q , donde p es el número de sílabas de la hembra entre coincidencias consecutivas y q es el número de sílabas del macho entre coincidencias consecutivas. Por ejemplo, el dúo de la Figura 6.1a muestra la secuencia de números $1/2, 1/2, 1/3, 1/3, 1/4$.

En la Figura 6.3a se muestra r_{ap} como función del tiempo promedio entre sílabas del macho T_{macho} para $n = 11$ dúos (de por lo menos 10 parejas diferentes). A fin de comparar entre dúos diferentes, el tiempo T_{macho} está normalizado al tiempo promedio entre sílabas de la hembra correspondiente, cuando está acoplada en $1/3$ (no todos los dúos recorren la secuencia completa $1/1-1/5$, pero todos pasan por el $1/3$ por lo menos una vez).

Una clara estructura tipo escalera surge de los datos, con escalones en los diferentes valores de r_{ap} encontrados. Notar que hay escalones con valores de p diferentes de 1, ubicados en la posición predecida por la *regla de suma de Farey* [González and Piro 1983]. Por ejemplo, encontramos un $2/7$, ubicado

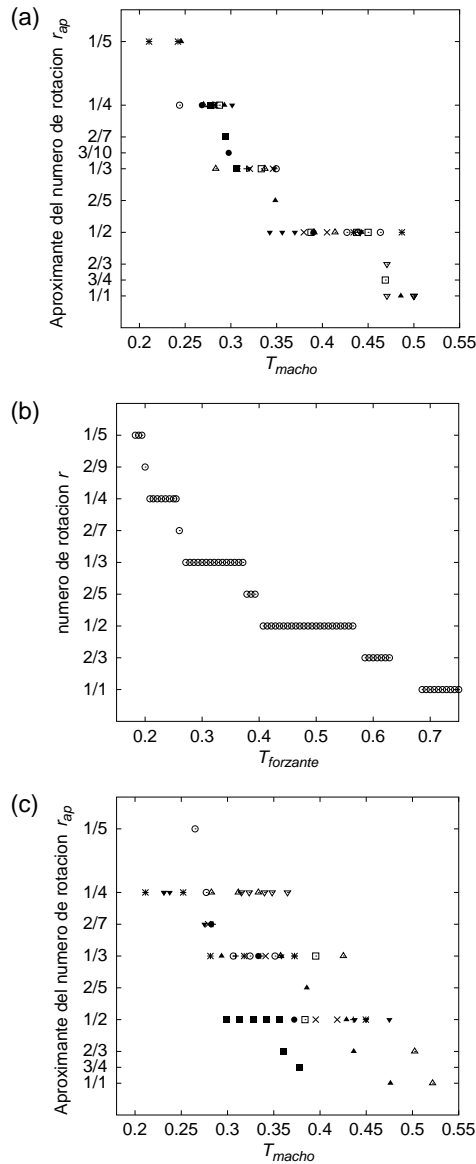


Figura 6.3: Acople de fase. (a) Aproximante del número de rotación r_{ap} como función del tiempo normalizado promedio entre sílabas del macho, para $n = 11$ dúos de hornero (de por lo menos 10 parejas diferentes). Símbolos diferentes corresponden a dúos diferentes. No todos los dúos realizaron la secuencia completa desde 1/1 hasta 1/5. (b) Número de rotación r como función del período T del forzante, para un oscilador no lineal forzado periódicamente. Esta estructura se conoce como “la escalera del diablo”. (c) Dúos al azar. La estructura de la escalera se pierde, lo que se refleja en el aumento del solapamiento entre escalones. Esto puede ser cuantificado definiendo el *solapamiento medio* ϕ entre escalones consecutivos (ver texto): los dúos originales tienen $\phi \sim 20\%$, mientras que los dúos al azar tienen $\phi \sim 70\%$.

entre los escalones $1/3$ y $1/4$. Un escalón más extraño todavía, el $3/10$, cayó entre los escalones $1/3$ y $2/7$.

Interpretación

Como veremos a continuación, esta estructura es característica de osciladores no lineales. Cuando un oscilador no lineal es forzado periódicamente muestra una diversidad de soluciones, tanto periódicas como no periódicas, que son huellas insoslayables de la naturaleza no lineal del sistema y que se organizan como muestra la Figura 6.3a. Nuestra interpretación nos dice que las sílabas del macho están actuando como un forzante unidireccional sobre el sistema de control motor vocal de la hembra, a través de un camino auditivo. Dado que las sílabas de la hembra conservan su identidad, los gestos respiratorio y siríngeo deben estar acoplados en algún nivel, formando un sistema no lineal tal que cuando es forzado presenta comportamiento subarmónico.

A continuación desarrollamos los conceptos que nos permiten sostener esta interpretación y explicar la estructura mostrada en la Figura 6.3a.

6.2 Horneros y osciladores forzados

6.2.1 Osciladores lineales vs. no lineales

Osciladores lineales y no lineales responden en forma cualitativamente distinta cuando son forzados periódicamente. Los osciladores lineales se acoplan a la frecuencia del forzante; es decir, siempre terminan siguiendo al forzante en un “uno a uno”. La amplitud del movimiento y la diferencia de fase con respecto al forzante pueden variar (dependiendo de la diferencia entre la frecuencia del forzante y la frecuencia propia del oscilador) pero, después de un transitorio, el oscilador lineal forzado termina vibrando con la misma frecuencia que el forzante.

Los osciladores no lineales, en cambio, responden con una variedad de comportamientos que dependen de la frecuencia y la amplitud del forzante. Si la frecuencia del forzante es similar a la frecuencia natural del oscilador forzado, el oscilador se acopla a la misma frecuencia que el forzante. Si la frecuencia del forzante es un poco mayor, las cosas no cambian demasiado: el oscilador sigue acoplándose en un “uno a uno”. Esto ocurre para un amplio rango de valores de frecuencia y amplitud del forzante. Sin embargo, si la frecuencia del forzante es suficientemente diferente de la frecuencia natural del oscilador, pueden ocurrir otros comportamientos, tanto periódicos como aperiódicos.

Por ejemplo, si la frecuencia del forzante es aproximadamente el doble de la frecuencia natural del oscilador no lineal, para ciertos valores de amplitud el oscilador forzado puede responder en “uno a dos”; es decir, cumpliendo sólo un ciclo cada dos ciclos del forzante. No es necesario que la frecuencia del forzante sea exactamente el doble de la frecuencia natural del oscilador; como en el caso del período 1, hay todo un rango de valores de parámetros para el cual el sistema se engancha en un período 2. En general, hay grandes regiones del espacio de parámetros (T, A) (período y amplitud del forzante) en las cuales el sistema presenta soluciones periódicas, repitiendo su comportamiento después de un número entero de veces el período del forzante. Es de destacar que en este

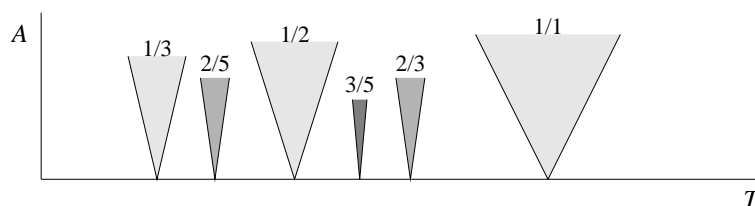


Figura 6.4: Lenguas de Arnold en sistemas no lineales forzados periódicamente. Para amplitudes bajas, la estructura de las lenguas de Arnold mostrada aquí es independiente de los detalles del sistema forzado o del forzante. Para amplitudes mayores, la geometría puede cambiar radicalmente de un sistema a otro (bifurcación de regiones, solapamiento, etc.).

tipo de comportamiento el oscilador vibra a una frecuencia que *no está presente* en el espectro del forzante.

6.2.2 Lenguas de Arnold y la escalera del diablo

Las regiones en el espacio de parámetros (T, A) donde ocurren soluciones periódicas con la misma periodicidad se conocen como *lenguas de Arnold* [Hoppensteadt and Izhikevich 1997]. Dicho comportamiento periódico puede ser descrito por el cociente de dos enteros $r = p/q$, llamado *número de rotación*, donde q es el período de la solución (medido en unidades del período del forzante) y p es el número de períodos que el oscilador forzado efectúa en q períodos del forzante.

Como está descrito en la subsección anterior, hay una gran región en el espacio (T, A) donde el oscilador presenta una solución de período 1. Esta solución se representa por el número de rotación $r = p_1/q_1 = 1/1$. Asimismo, hay otra región en el espacio de parámetros donde ocurre una solución de período 2, representada por $r = p_2/q_2 = 1/2$. La región más grande entre estas dos regiones que presenta soluciones periódicas es una en la cual hay una solución de período 3, caracterizada por

$$r = \frac{p_1 + p_2}{q_1 + q_2} = \frac{2}{3} \quad (6.1)$$

Esta es la “regla de suma de Farey”, y se cumple para un gran conjunto de sistemas no lineales. El sistema forzado puede ser un oscilador de relajación o un sistema excitable; el forzante periódico puede ser una función armónica o una secuencia de pulsos. Lo importante es que el comportamiento subarmónico recién descrito se debe únicamente a la naturaleza no lineal del sistema forzado, y no a sus detalles “internos”. Un esquema parcial de la estructura de las lenguas de Arnold (para valores chicos de amplitud) se muestra en la Figura 6.4.

Un corte unidimensional al espacio de parámetros (T, A) nos da bastante información sobre la dinámica del sistema. Si recorremos el espacio (T, A) a lo largo de una recta horizontal (o sea, A constante), podemos graficar los números de rotación obtenidos para cada período T . En la Figura 6.3b se muestra el resultado de esta operación para un oscilador de relajación integrado numéricamente. Una estructura de escalera surge de los datos, como la descrita en los dúos de horneros de la sección anterior.

6.2.3 Dúos de prueba

Como una prueba de la plausibilidad de nuestra hipótesis, utilizamos nuestros modelos de fonación para generar dúos artificiales. El modelo, desarrollado en la Sección 3.3.4, es

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\epsilon x + py - cx^2y,\end{aligned}\tag{6.2}$$

donde ϵ es la constante de restitución elástica por unidad de masa del labium y p es una función lineal de la presión del saco aéreo. Este modelo fue originalmente diseñado para simular vocalizaciones de Oscinos, cuya siringe muestra considerables diferencias con la de los horneros. Una de las desventajas más notorias es que la frecuencia de la vocalización en Suboscinos podría estar correlacionada con la presión del saco aéreo, como ocurre en palomas [Beckers et al. 2003b]. Utilizamos este modelo simplemente como una cruda aproximación a la fonación en horneros.

En el espíritu del capítulo 4, para obtener las sílabas del macho integramos la ecuación (6.2) tomando ϵ y p como funciones armónicas del tiempo, pero con frecuencia creciente entre 5 y 20 Hz. Esto es:

$$\begin{aligned}\epsilon(t) &= \epsilon_0 + \epsilon_1 f(t) \\ p(t) &= p_0 + p_1 g(t),\end{aligned}\tag{6.3}$$

donde $f(t) = \cos(\omega t)$ y $g(t) = \cos(\omega t + \phi_0)$, con ω creciente en el tiempo. Llamamos a $\epsilon(t)$ y $p(t)$ los *gestos motores* del macho.

Análogamente, generamos las sílabas de la hembra con las ecuaciones (6.2) y (6.3) pero en este caso las funciones $f(t)$ y $g(t)$ constituyen un oscilador no lineal forzado a su vez por los gestos del macho. Para una mayor similitud con las sílabas reales, utilizamos $f(t) = u - v$ y $g(t) = u + v$, con u y v dados por

$$\begin{aligned}\dot{u} &= 75(u - u^3 + 0.5 \cos(\omega(t)t) + 0.5) \\ \dot{v} &= 6(u + 0.7 - 0.8v).\end{aligned}\tag{6.4}$$

Es conveniente aclarar que no estamos suponiendo nada sobre el sustrato neuronal responsable de crear los gestos vocales de la hembra. Lo que sí estamos suponiendo es que dicho circuito tiene que estar afectado por las señales auditivas que le llegan del macho, y que los gestos vocales de la hembra pueden ser modelados finalmente por un oscilador de relajación muy simple (o, al menos, que podemos modelar la forma en que reaccionan ante un forzante). El resultado se muestra en la Figura 6.1b, donde se ven acoples a números de rotación $r_{ap} = 1/2, 1/2, 1/2, 2/5, 1/3, 1/3$ y $1/4$.

La estructura mostrada en la Figura 6.3a es muy sugestiva, pero alguien podría decir: “Puede ser una casualidad, debida al desfase de dos frecuencias que varían independientemente”. Para descartar esta hipótesis, construimos dúos al azar en base a los datos originales y los analizamos con el mismo protocolo. La construcción de los dúos al azar fue como sigue. Primero elegimos aleatoriamente dos dúos originales. Luego, tomamos los sonogramas correspondientes y los superpusimos, desplazando uno respecto del otro tal que las primeras sílabas de los machos coincidieran. Por último, las sílabas del macho del

primer dúo y de la hembra del segundo dúo fueron borrados del sonograma. De este modo, obtuvimos 11 nuevos sonogramas de dúos formados al azar, y los sometimos al mismo análisis que los dúos originales. El resultado se ve en la Figura 6.3c, donde se aprecia que la estructura de la escalera está rota. Como modo de cuantificar una mayor o menor apariencia de escalera, definimos el solapamiento medio entre escalones consecutivos ϕ , usando los escalones 1/2, 1/3 y 1/4 (los únicos escalones para los que podemos estimar su largo). Primero se toman los escalones 1/2 y 1/3 y se calcula el cociente entre longitud solapada y longitud promedio (distancias medidas en la dirección horizontal). Lo mismo se hace con los escalones 1/3 y 1/4. Los dos cocientes obtenidos se promedian para obtener finalmente ϕ . En el caso de los dúos al azar (Figura 6.3c), el solapamiento medio es $\phi \sim 70\%$, mientras que en los dúos originales es $\phi \sim 20\%$ (Figura 6.3a).

6.2.4 El oscilador en Horneros

Los ritmos corporales en organismos vivos son ubicuos en la Naturaleza: desde el paradigmático latido del corazón hasta los ciclos circadianos, pasando por el temblor parkinsoniano y el ritmo respiratorio, los ritmos biológicos reflejan siempre la existencia de un sistema dinámico no lineal subyacente [Glass 2001].

La no linealidad es una característica necesaria en los osciladores biológicos; de este modo, por ejemplo, la frecuencia y la amplitud de la oscilación no se ven afectadas por perturbaciones transitorias [Glass 2001]. En el mismo sentido, dado que los osciladores biológicos son osciladores no lineales estables, muchas de las características generales de la interacción entre el forzante periódico y el ritmo del oscilador forzado pueden ser predichas sin conocer el detalle del oscilador o del forzante.

Una pregunta fundamental con respecto a osciladores biológicos se refiere a la ubicación anatómica del oscilador; es decir, cuál es el sustrato neuronal que le da origen. En Oscinos, se cree que el “reloj” que da el ritmo básico del canto se encuentra en el núcleo NIf [Farries 2001, Farries 2004], que recibe información auditiva y enerva al HVc (ver Sección 4.3). Por otro lado, la interacción entre la circuitería interna del RA y las instrucciones que el HVc manda al RA pueden ser fundamentales para marcar el ritmo del canto, como sugieren las soluciones subarmónicas mostradas en [Abarbanel et al. 2003].

El estudio del cerebro de los no-Oscinos reveló la existencia de un “camino de control motor general”, que se asemeja al camino de control motor del canto de los Oscinos en su organización anatómica [Farries 2001, Farries 2004]. Aunque no se han encontrado núcleos telencefálicos especializados como el HVc y el RA en el cerebro de los no-Oscinos,² se cree que todas las aves comparten el mismo sistema vocal-respiratorio de núcleos medulares nXIIits-PAm/RAm [Farries 2001].

El circuito neuronal respiratorio es el que provee las instrucciones neuronales para la espiración y, por lo tanto, marca el ritmo del canto. Tanto en Oscinos como en no-Oscinos, los sacos aéreos son gobernados por músculos inspiratorios y expiratorios; por medio de neuronas motoras, estos músculos son gobernados por los núcleos medulares PAm y RAm, respectivamente, que están conectados

²Con excepción de los loros y los colibríes. Estos dos grupos tienen siete núcleos neuronales telencefálicos involucrados en el aprendizaje y producción vocal muy similares a los de los Oscinos [Jarvis et al. 2000].

entre sí y se cree que son mutuamente inhibitorios [Keener and Sneyd 1998]. A partir de aquí comienzan las diferencias más sustanciales: los no-Oscinos carecen de núcleos telencefálicos especializados en el canto. Los núcleos medulares nXIIIts y PAm/RAm son enervados, en cambio, por el núcleo DM; se cree que el DM es capaz de coordinar la producción de vocalizaciones simples [Farries 2001]. Con respecto a la influencia del circuito auditivo sobre el circuito respiratorio-vocal, es poco lo que se sabe; el núcleo DM se encuentra adyacente al sistema auditivo, pero todavía no se conocen conexiones anatómicas entre los centros auditivos y vocales a nivel de la médula [Wild 1997].

Últimas noticias

Un reciente trabajo [Amador et al. 2005] modeló la actividad de los núcleos respiratorios descritos recién para explicar el origen de la subarmonicidad en el dúo de horneros. Asumen que el DM recibe entradas desde el sistema auditivo. Los núcleos PAm/RAm reciben instrucciones del DM, y envían señales que se traducen en instrucciones motoras para los músculos respiratorios. El modelo incorpora además la actividad de los sensores de CO₂, que tienen un efecto inhibitorio sobre el núcleo de la inspiración PAm. Encontraron que una variedad de estructuras rítmicas subarmónicas pueden ser generadas por un sustrato neuronal simple en la hembra del hornero, con una organización compatible con la encontrada experimentalmente [Amador et al. 2005]. La realimentación que llega al PAm a través de los sensores de CO₂ cumple un rol fundamental en el circuito, aportando evidencia indirecta para entender el comportamiento como resultado de la interacción neural-corporal, en lugar de la relación amo-esclavo tradicionalmente considerada.

Capítulo 7

Conclusiones

Ocurre que en general no se sopesa suficientemente el fundamento biológico del lenguaje cuando se lo describe como un factor más entre los constituyentes de la cultura. El lenguaje es la articulación más importante, misteriosa e impenetrable entre cuerpo y sociedad. Resulta una humillante estimulación para la inteligencia científica que aquello que nos diferencia de todas las otras especies animales se encuentre tan remotamente alejado de nuestra comprensión, en particular en lo que se refiere a su nacimiento y evolución.

Ivonne Bordelois, *La palabra amenazada* (2003).

Hablar de lenguaje en el marco de una tesis de física podría parecer una extravagancia. Sin embargo, la distancia entre el párrafo de Bordelois y el tema tratado en esta tesis es elocuente: tal es la distancia que nos separa de los insondables —por ahora— mecanismos del cerebro humano.

Paradójicamente, la manifiesta lejanía entre Bordelois y el sistema vocal de las aves señala al mismo tiempo la importancia del canto de aves como modelo para entender nuestro cerebro y alguna de sus expresiones; no ya el lenguaje sino simplemente el habla, y aún el aprendizaje sensorial y motor en general.

El aprendizaje vocal es un rasgo poco común en la naturaleza. Hasta donde se sabe, el aprendizaje vocal es poseído sólo por tres grupos distantes de mamíferos (humanos, cetáceos y murciélagos), y por tres grupos distantes de aves (aves canoras, loros y colibríes) [Jarvis 2004, Brainard and Doupe 2002]. ¿Qué hay de especial en el cerebro de estos animales que imitan sonidos? En este sentido, el aprendizaje del canto en aves tiene tantas similitudes con la adquisición del habla en humanos [Brainard and Doupe 2002] y tantas ventajas experimentales [Brenowitz et al. 1997] que lo convierten en un modelo invaluable de aprendizaje vocal.

El canto es el resultado de la traducción de un programa central de control motor a la acción coordinada de los sistemas motores periféricos: los grupos

musculares siríngicos y respiratorios [Suthers 1997]. Entender cómo el cerebro organiza este complejo comportamiento es de central importancia [Fee et al. 1998]. Una gran cantidad de trabajo está dedicado actualmente a desentrañar la representación cerebral del canto y los mecanismos por medio de los cuales dicha representación es modificada por la experiencia [Brainard and Doupe 2002]. Estos acercamientos “desde arriba” son desafiados siempre por la abrumadora cantidad de datos que surgen de los registros de actividad neuronal central, aún en el caso de las aves donde la interpretación es en principio más accesible dada la alta especialización de los centros neuronales involucrados [Farries 2001].

Además de la complejidad del control neuronal, el investigador debe enfrentar los aspectos complejos del comportamiento debidos pura y exclusivamente a la dinámica inherente al aparato vocal: la no linealidad intrínseca del aparato vocal puede provocar que una variación leve en las instrucciones neuronales sea traducida en cambios abruptos en el sonido emitido [Fee et al. 1998]. Es necesario por lo tanto un conocimiento profundo del funcionamiento del aparato vocal, el dispositivo mecánico que media entre el sustrato neuronal y el comportamiento observado. Este acercamiento “desde abajo” es novedoso y puede aportar un nuevo punto de vista en el entendimiento del control motor [Suthers and Margoliash 2002].

Los mecanismos físicos de la fonación en aves fueron objeto de especulación durante mucho tiempo, tanto como la función de diferentes estructuras presentes en la siringe. Estas especulaciones tuvieron su fin con los trabajos de Goller, Suthers y Larsen [Goller and Suthers 1996a, Goller and Suthers 1996b, Goller and Larsen 1997a, Goller and Larsen 1997b], quienes determinaron con precisión el origen del sonido y el rol de cada músculo en la siringe. En contraposición, el estudio de la voz humana cuenta con una gran variedad de modelos basados en observaciones experimentales. En particular, destacamos el sencillo modelo flameante de Titze [Titze 1988].

En esta tesis analizamos y modelamos el comportamiento del aparato vocal de las aves canoras con el espíritu y las herramientas de la Dinámica No Lineal. Desarrollamos un modelo de fonación muy simple [Laje et al. 2002], basado en los trabajos teóricos sobre la voz humana de Titze y en las detalladas mediciones de Goller, Suthers y Larsen. Los parámetros de este modelo tienen estrecha relación con datos fisiológicos relevantes, como presión del saco aéreo y actividad muscular en la siringe. Las características cualitativas del diagrama de bifurcaciones de este modelo nos ofrecieron predicciones concretas sobre actividad muscular que fueron confirmadas experimentalmente [Mindlin et al. 2003].

Un aspecto central de la Dinámica No Lineal es la búsqueda de modelos paradigmáticos que tengan un comportamiento deseado, como el oscilador de relajación que utilizamos para modelar la siringe. Gracias a esta característica, nuestro trabajo nos permitió aventurar un paso hacia arriba: desde la física del aparato fonador hacia el cerebro del ave. Construyendo sobre preguntas físicas concretas, nuestros resultados nos concedieron una intuición sobre los mecanismos periféricos del control motor del canto. Propusimos que una gran variedad de vocalizaciones pueden ser interpretadas en términos de dos gestos motores básicos: la presión del saco aéreo y la tensión de los labios [Gardner et al. 2001]. La variedad de vocalizaciones posibles crece con el agregado de un tercer gesto motor, el control activo del flujo de aire [Laje et al. 2002]. Estos gestos motores están bien identificados experimentalmente, son más o menos estereotipados y tienen interpretación directa en función de los parámetros de nuestro modelo.

Una consecuencia de este resultado es que sería necesario relacionar la actividad motora central con los gestos motores propuestos, más que con las propiedades acústicas del sonido emitido [Suthers and Margoliash 2002].

Sabemos que la no linealidad intrínseca de la siringe puede aportar una cuota de complejidad al sonido emitido, independientemente de la ocasional complejidad de las instrucciones neuronales. Esta no linealidad intrínseca se manifiesta en el sonido emitido como frecuencias subarmónicas, duplicaciones de período, saltos de frecuencia, etc. En esta tesis mostramos que un posible origen dinámico de dichos sonidos complejos es la realimentación retrasada que ocurre cuando fuente y tracto interactúan acústicamente, y propusimos experimentos sencillos para decidir entre causas posibles de sonidos complejos [Laje and Mindlin 2005, Laje et al. 2001]. Mostramos además que la interacción acústica es también un posible origen de interacción entre las dos fuentes sonoras de los Oscinos [Sciamarella et al. 2005].

Finalmente, nuestro trabajo sobre gestos motores nos dió un marco para interpretar la llamativa organización subyacente en el canto a dúo de los horneros [Laje and Mindlin 2003]. Macho y hembra dentro de un dúo alternan sílabas en forma aparentemente descoordinada; sin embargo, encontramos que las sílabas están organizadas según una regla subyacente no trivial, compatible con la organización de acople de fase de los osciladores no lineales forzados. Nuestra hipótesis es que los centros neuronales que gobiernan los gestos motores de la hembra hornero reaccionan a la información auditiva del canto del macho como un oscilador no lineal forzado periódicamente. Esta es una de las pocas evidencias de sincronización de ritmos corporales en animales intactos.

Apéndice A

Detalles de los modelos fuente-tracto y fuente-fuente

En este apéndice damos los detalles del desarrollo de los modelos de acople acústico presentados en el Capítulo 5.

A.1 Acople acústico entre fuente y tracto

En esta sección obtendremos la expresión para $s(t)$ dada por la ecuación (5.13). Comenzamos considerando una onda acústica armónica $s(t)$ que diverge localmente de la válvula vocal con un frente esférico. Esta es una onda de la forma

$$s(r, t) = \frac{P_0}{r} e^{i(\omega t - kr)}. \quad (\text{A.1})$$

El número (complejo) P_0 se determina a través de las condiciones de contorno, que son establecidas en la fuente (esto es, justo a la salida de la válvula vocal donde podemos conocer la velocidad de partícula $v(t)$). Para establecer las condiciones de contorno y determinar P_0 , usamos la relación entre s y v para una onda esférica, a una distancia d_1 de la fuente. Tomamos d_1 mucho más chica que la longitud de onda λ , pero comparable al tamaño de la fuente h , de acuerdo con nuestra hipótesis local. De esta manera,

$$s(d_1, t) = Z(d_1) v(d_1, t), \quad (\text{A.2})$$

donde $Z = R + i\omega I$ es la impedancia acústica específica [Kinsler et al. 1982] (redefinida por unidad de M). R e I son las resistencia e inercia acústicas específicas, respectivamente, y están dadas por

$$R = \frac{\rho c}{M} \frac{(kd_1)^2}{1 + (kd_1)^2} \quad (\text{A.3})$$

$$I = \frac{\rho}{M} \frac{d_1}{1 + (kd_1)^2}, \quad (\text{A.4})$$

donde ρ es la densidad del aire y c la velocidad del sonido [Kinsler et al. 1982]. R e I dependen de la frecuencia del sonido f a través del número de onda $k = 2\pi f/c$. La constante P_0 puede calcularse al evaluar la ecuación (A.1) en $r = d_1$ e igualando a la ecuación (A.2). Con esto obtenemos

$$P_0 = e^{-i(\omega t - kd_1)} d_1 Z(d_1) v(d_1, t), \quad (\text{A.5})$$

que no depende del tiempo, dado que la velocidad de partícula también es armónica. Evaluando la ecuación (A.1) a una distancia d_2 de la fuente (esto es, en la base de la tráquea donde vale la superposición de ondas planas, ver Figura 5.3a), obtenemos una expresión para la perturbación de la presión $s(t)$ a una distancia d_2 en términos de la velocidad de partícula $v(t)$ en la fuente:

$$s(d_2, t) = \frac{d_1}{d_2} Z(d_1) e^{-ik(d_2 - d_1)} v(d_1, t). \quad (\text{A.6})$$

Para una onda acústica armónica, la velocidad de partícula también es armónica. Por lo tanto podemos escribir $\dot{v} = i\omega v$ y reescribir la última expresión como

$$s(d_2, t) = \mathcal{R} v(d_1, t) + \mathcal{I} \dot{v}(d_1, t), \quad (\text{A.7})$$

donde definimos

$$\mathcal{R} = \frac{d_1}{d_2} [R \cos k(d_2 - d_1) + \omega I \sin k(d_2 - d_1)] \quad (\text{A.8})$$

$$\mathcal{I} = \frac{d_1}{d_2} [I \cos k(d_2 - d_1) - \frac{R}{\omega} \sin k(d_2 - d_1)]. \quad (\text{A.9})$$

Notemos en la ecuación (A.7) que en general hay una diferencia de fase entre presión s y velocidad v , gracias al término con \dot{v} . Este resultado nos dice que la diferencia de fase entre presión acústica y velocidad de partícula se debe no sólo a la geometría esférica de la onda sino también a que presión y velocidad están evaluadas en diferentes posiciones. Teniendo esto en cuenta, quitamos d_1 y d_2 de los argumentos de $v(d_1, t)$ y $s(d_2, t)$ desde ahora en adelante y escribimos simplemente $v(d_1, t) \equiv v(t)$ and $s(d_2, t) \equiv s(t)$. Notar que si el final de la región de propagación esférica está suficientemente cercano a la región de propagación plana (o sea $d_2 - d_1 \ll \lambda$) podemos escribir simplemente $\mathcal{R} \sim R$ y $\mathcal{I} \sim I$. Con esto, que se asume de ahora en adelante, la ecuación (A.7) se transforma en

$$s(t) = Rv(t) + I\dot{v}(t), \quad (\text{A.10})$$

que es idéntica a la ecuación (5.9).

Para terminar, falta una expresión de $v(t)$ en términos de las variables del modelo x e y , para relacionar la presión en el tracto con la dinámica de los labia. Consistentemente con nuestra hipótesis previa (flujo divergente esférico cerca de la fuente), asumimos que el flujo a través de la válvula (que tiene área de salida a_2 , ver Sección 3.3.3) adquiere un perfil esférico con área A_s a una distancia d_1 . Sin embargo, la válvula no radia en todas direcciones, sino en una fracción del frente esférico. Para tener en cuenta esto escribimos $A_s = 4\pi d_1^2 q$, con $q < 1$. La conservación del flujo en la pequeña región (de tamaño $d_1 \ll \lambda$) entre a_2 y A_s nos permite escribir

$$v(t) = \frac{a_2(t)}{A_s} v_0 = \frac{2hv_0}{A_s} (x_0 + x - \tau y), \quad (\text{A.11})$$

donde v_0 es la velocidad media de partícula dada por Bernouilli: $v_0 = \sqrt{2p_s M/\rho}$. Finalmente reemplazamos la ecuación (A.11) en la ecuación (A.10), tiramos el término constante (que no aporta a la presión acústica) y redefinimos los coeficientes, para obtener

$$s(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}), \quad (\text{A.12})$$

que es idéntica a la ecuación (5.13). Los coeficientes de acople son

$$\alpha = 2h \frac{v_0}{A_s} R \quad (\text{A.13})$$

$$\beta = 2h \frac{v_0}{A_s} I. \quad (\text{A.14})$$

A.2 Modelo de una fuente

Aquí daremos los detalles del modelo presentado en la Sección 5.2.2.

A.2.1 Forma explícita

El modelo completo de una fuente acoplada acústicamente al tracto vocal con realimentación retrasada se arma con las ecuaciones (5.3), (5.4), (5.17) y (5.18):

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= f(x, y) + p_g \end{aligned} \quad (\text{A.15})$$

$$f(x, y) = -\epsilon x - by - cx^2 y - f_0 \quad (\text{A.16})$$

$$p_g = p_s + (Dy - 1)(p_s - p_i) \quad (\text{A.17})$$

$$p_i(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}) - \gamma p_i(t - T), \quad (\text{A.18})$$

Nótese que la ecuación (A.18) introduce un término de aceleración \dot{y} en el miembro derecho de la ley de Newton, ecuación (A.15). Dicho término es heredado directamente del modelo flameante y refleja el hecho de que el flujo que sale de la válvula no está en fase con x (ver Sección 3.3.3). Este término hace que el sistema pierda su forma explícita $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ e impide el uso de un integrador numérico estándar tipo Runge-Kutta. Además, involucra una ecuación algebraica con retraso. Sin embargo, el sistema puede ser reescrito en forma apropiada para integrarlo con un Runge-Kutta.

Para poner el sistema en forma explícita, primero separamos el término de aceleración en la ecuación (A.18) y escribimos

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= f(x, y) + \tilde{p}_g + (Dy - 1)\beta\tau\dot{y} \end{aligned} \quad (\text{A.19})$$

$$f(x, y) = -\epsilon x - by - cx^2 y - f_0 \quad (\text{A.20})$$

$$\tilde{p}_g = p_s + (Dy - 1)(p_s - \tilde{p}_i) \quad (\text{A.21})$$

$$\tilde{p}_i(t) = \alpha(x - \tau y) + \beta y - \gamma p_i(t - T) \quad (\text{A.22})$$

$$p_i(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}) - \gamma p_i(t - T). \quad (\text{A.23})$$

Notar la definición de \tilde{p}_g y especialmente la de \tilde{p}_i (en contraste con las de p_g y p_i). Ahora despejamos \dot{y} en la ecuación (A.19):

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= [f(x, y) + \tilde{p}_g] [1 - (Dy - 1)\beta\tau]^{-1}\end{aligned}\quad (\text{A.24})$$

$$f(x, y) = -\epsilon x - by - cx^2y - f_0 \quad (\text{A.25})$$

$$\tilde{p}_g = p_s + (Dy - 1)(p_s - \tilde{p}_i) \quad (\text{A.26})$$

$$\tilde{p}_i(t) = \alpha(x - \tau y) + \beta y - \gamma p_i(t - T) \quad (\text{A.27})$$

$$p_i(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}) - \gamma p_i(t - T), \quad (\text{A.28})$$

con lo que el sistema queda en forma explícita.

Notar la diferencia entre los argumentos de p_i en la ecuación (A.27) y en la ecuación (A.28). Esta diferencia nos permite resolver numéricamente el sistema como sigue: primero integre la ecuación (A.24) en el instante t junto con las definiciones (A.25), (A.26) y (A.27), para una dada $p_i(t - T)$ (ya sea como condición inicial o calculada previamente); después calcule p_i con la ecuación (A.28) para ser utilizada más tarde en el instante $t + T$. Esto puede hacerse llevando cuenta de los valores de p_i en cada paso temporal desde $t - T$ hasta t (un intervalo temporal de longitud T antes de t). Las condiciones iniciales para este sistema son $x(t = 0) = x_0$, $y(t = 0) = y_0$ y $p_i(t \in [-T, 0]) = p_{i0}(t)$. Notar que p_i debe estar inicialmente definida en un intervalo de longitud T .

A.2.2 Aproximación de bajo acople

Vale la pena estudiar las condiciones bajo las cuales es posible despreciar el término $(Dy - A)\beta\tau$ frente a 1 en el denominador de la ecuación (A.24).

En general hay más de una manera de adimensionalizar una ecuación. Para hacer la aproximación correcta procederemos como en [Strogatz 1994]. Notemos primero que el término Dy debe compararse a 1. Esto, junto con la adimensionalización de las variables x y t y de las mismas ecuaciones que hace que todas las variables y las derivadas sean de orden 1, nos lleva a que

$$(Dy - 1)\beta\tau \ll 1 \quad \text{si} \quad \beta \ll \tau^{-1} \quad (\text{A.29})$$

Esta aproximación puede interpretarse de varias formas. La más directa es la de bajo acople: la aproximación es válida sólo si β es suficientemente pequeño, lo que a frecuencia fija depende de parámetros acústicos y anatómicos. Una segunda interpretación está relacionada con la diferencia de fase entre las porciones superior e inferior del labium en el movimiento de flameo, que es directamente establecida por el valor de τ (ecuaciones 3.2 y 3.3). Desde este punto de vista, la condición es satisfecha para valores pequeños de τ , o sea pequeña diferencia de fase. La tercera interpretación se basa en el comportamiento de β como función de la frecuencia. Como β crece con la frecuencia (Figura 5.3b), la condición se satisface para valores de frecuencia suficientemente grandes.

Con esta aproximación, podemos reemplazar el denominador de la ecuación (A.24) por 1, con lo cual el sistema queda finalmente

$$\dot{x} = y$$

$$\dot{y} = f(x, y) + \tilde{p}_g \quad (\text{A.30})$$

$$f(x, y) = -\epsilon x - by - cx^2y - f_0 \quad (\text{A.31})$$

$$\tilde{p}_g = p_s + (Dy - 1)(p_s - \tilde{p}_i) \quad (\text{A.32})$$

$$\tilde{p}_i(t) = \alpha(x - \tau y) + \beta y - \gamma p_i(t - T) \quad (\text{A.33})$$

$$p_i(t) = \alpha(x - \tau y) + \beta(y - \tau \dot{y}) - \gamma p_i(t - T), \quad (\text{A.34})$$

que es idéntico a las ecuaciones (5.22)–(5.25).

A.3 Modelo de dos fuentes

Damos aquí los detalles del modelo presentado en la Sección 5.3.2.

A.3.1 Forma explícita

Identifiquemos las fuentes con l (izquierda) y r (derecha). El modelo para dos fuentes acopladas al tracto se forma con las ecuaciones (5.3), (5.4), (5.17) y (5.18):

$$\begin{aligned} \dot{x}_l &= y_l \\ \dot{y}_l &= f_l(x_l, y_l) + p_{gl} \end{aligned} \quad (\text{A.35})$$

$$\begin{aligned} \dot{x}_r &= y_r \\ \dot{y}_r &= f_r(x_r, y_r) + p_{gr} \end{aligned} \quad (\text{A.36})$$

$$f_j(x_j, y_j) = -\epsilon_j x_j - b_j y_j - c_j x_j^2 y_j - f_{0j} \quad (\text{A.37})$$

$$p_{gj} = p_s + (D_j y_j - 1)(p_s - p_i) \quad (\text{A.38})$$

$$\begin{aligned} p_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l(y_l - \tau_l \dot{y}_l) \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r(y_r - \tau_r \dot{y}_r) - \gamma p_i(t - T) \end{aligned} \quad (\text{A.39})$$

($j = l, r$). Notar que p_i se forma con la contribución de las dos fuentes, más la onda reflejada. La presión p_i es común a ambas fuentes (dado que está definida en la base de la tr'aquea), como también lo es la presión sublabial p_s que está asociada a la presión del saco aéreo.

Siguiendo pasos análogos a los de la Sección A.2, separamos los términos de aceleración en la ecuación (A.39) y escribimos

$$\begin{aligned} \dot{x}_l &= y_l \\ \dot{y}_l &= f_l(x_l, y_l) + \tilde{p}_{gl} \\ &\quad + (D_l y_l - 1)(\beta_l \tau_l \dot{y}_l + \beta_r \tau_r \dot{y}_r) \end{aligned} \quad (\text{A.40})$$

$$\begin{aligned} \dot{x}_r &= y_r \\ \dot{y}_r &= f_r(x_r, y_r) + \tilde{p}_{gr} \\ &\quad + (D_r y_r - 1)(\beta_l \tau_l \dot{y}_l + \beta_r \tau_r \dot{y}_r) \end{aligned} \quad (\text{A.41})$$

$$f_j(x_j, y_j) = -\epsilon_j x_j - b_j y_j - c_j x_j^2 y_j - f_{0j} \quad (\text{A.42})$$

$$\tilde{p}_{gj} = p_s + (D_j y_j - 1)(p_s - \tilde{p}_i) \quad (\text{A.43})$$

$$\begin{aligned}\tilde{p}_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l y_l \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r y_r - \gamma p_i(t - T)\end{aligned}\quad (\text{A.44})$$

$$\begin{aligned}p_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l(y_l - \tau_l \dot{y}_l) \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r(y_r - \tau_r \dot{y}_r) - \gamma p_i(t - T).\end{aligned}\quad (\text{A.45})$$

Notar otra vez la definición de \tilde{p}_{gj} y \tilde{p}_i , análogas a las de la sección anterior. Ahora las ecuaciones (A.40) y (A.41) son dos ecuaciones acopladas lineales en \dot{y}_l y \dot{y}_r . Despejando \dot{y}_r en ecuación (A.41) y reemplazando en la ecuación (A.40), obtenemos una expresión para \dot{y}_l :

$$\begin{aligned}\dot{y}_l &= \{[1 - (D_r y_r - 1)\beta_r \tau_r][f_l(x_l, y_l) + \tilde{p}_{gl}] \\ &\quad + (D_l y_l - 1)\beta_r \tau_r [f_r(x_r, y_r) + \tilde{p}_{gr}]\} \\ &\quad \times [1 - (D_l y_l - 1)\beta_l \tau_l \\ &\quad - (D_r y_r - 1)\beta_r \tau_r]^{-1}.\end{aligned}\quad (\text{A.46})$$

Una expresión similar (intercambiando los subíndices $l \leftrightarrow r$) resulta para \dot{y}_r . Con esto, podemos poner el sistema en forma explícita:

$$\begin{aligned}\dot{x}_l &= y_l \\ \dot{y}_l &= \{[1 - (D_r y_r - 1)\beta_r \tau_r][f_l(x_l, y_l) + \tilde{p}_{gl}] \\ &\quad + (D_l y_l - 1)\beta_r \tau_r [f_r(x_r, y_r) + \tilde{p}_{gr}]\} \\ &\quad \times [1 - (D_l y_l - 1)\beta_l \tau_l \\ &\quad - (D_r y_r - 1)\beta_r \tau_r]^{-1}\end{aligned}\quad (\text{A.47})$$

$$\begin{aligned}\dot{x}_r &= y_r \\ \dot{y}_r &= \{[1 - (D_l y_l - 1)\beta_l \tau_l][f_r(x_r, y_r) + \tilde{p}_{gr}] \\ &\quad + (D_r y_r - 1)\beta_l \tau_l [f_l(x_l, y_l) + \tilde{p}_{gl}]\} \\ &\quad \times [1 - (D_l y_l - 1)\beta_l \tau_l \\ &\quad - (D_r y_r - 1)\beta_r \tau_r]^{-1}\end{aligned}\quad (\text{A.48})$$

$$f_j(x_j, y_j) = -\epsilon_j x_j - b_j y_j - c_j x_j^2 y_j - f_{0j}\quad (\text{A.49})$$

$$\tilde{p}_{gj} = p_s + (D_j y_j - 1)(p_s - \tilde{p}_i)\quad (\text{A.50})$$

$$\begin{aligned}\tilde{p}_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l y_l \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r y_r - \gamma p_i(t - T)\end{aligned}\quad (\text{A.51})$$

$$\begin{aligned}p_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l(y_l - \tau_l \dot{y}_l) \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r(y_r - \tau_r \dot{y}_r) - \gamma p_i(t - T).\end{aligned}\quad (\text{A.52})$$

Nótese nuevamente la diferencia entre los argumentos de p_i en la ecuación (A.51) y en la ecuación (A.52). Las observaciones hechas en la sección anterior en relación a la integración numérica del sistema son válidas aquí.

A.3.2 Aproximación de bajo acople

La misma aproximación realizada en la Sección A.2 puede realizarse aquí, a fin de obtener una forma mucho más simple para nuestro modelo. Obtenemos que

$$(D_j y_j - 1)\beta_j \tau_j \ll 1 \quad \text{si} \quad \beta_j \ll \tau_j^{-1}\quad (\text{A.53})$$

($j = l, r$). Notar que esta condición nos permite aproximar los denominadores de las ecuaciones (A.47) y (A.48) por 1, como antes, y además nos permite despreciar los términos cruzados en sus numeradores. Específicamente, en el numerador de la ecuación (A.47) por ejemplo, nos referimos a

$$\begin{aligned} (D_r y_r - 1)\beta_r \tau_r &\ll 1 \\ (D_l y_l - 1)\beta_r \tau_r [f_r(x_r, y_r) + \tilde{p}_{gr}] &\ll f_l(x_l, y_l) + \tilde{p}_{gl}. \end{aligned}$$

El sistema aproximado es, finalmente,

$$\begin{aligned} \dot{x}_l &= y_l \\ \dot{y}_l &= f_l(x_l, y_l) + \tilde{p}_{gl} \end{aligned} \quad (\text{A.54})$$

$$\begin{aligned} \dot{x}_r &= y_r \\ \dot{y}_r &= f_r(x_r, y_r) + \tilde{p}_{gr} \end{aligned} \quad (\text{A.55})$$

$$f_j(x_j, y_j) = -\epsilon_j x_j - b_j y_j - c_j x_j^2 y_j - f_{0j} \quad (\text{A.56})$$

$$\tilde{p}_{gj} = p_s + (D_j y_j - 1)(p_s - \tilde{p}_i) \quad (\text{A.57})$$

$$\begin{aligned} \tilde{p}_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l y_l \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r y_r - \gamma p_i(t - T) \end{aligned} \quad (\text{A.58})$$

$$\begin{aligned} p_i(t) &= \alpha_l(x_l - \tau_l y_l) + \beta_l(y_l - \tau_l \dot{y}_l) \\ &\quad + \alpha_r(x_r - \tau_r y_r) + \beta_r(y_r - \tau_r \dot{y}_r) - \gamma p_i(t - T) \end{aligned} \quad (\text{A.59})$$

($j = l, r$), que es idéntico a las ecuaciones (5.40)–(5.45).

ϵ	$= 3 \times 10^8$	s^{-2}
b	$= 2 \times 10^4$	s^{-1}
c	$= 2 \times 10^8$	$\text{s}^{-1} \text{cm}^{-2}$
f_0	$= 4.9 \times 10^4$	cm s^{-2}
p_s	$= 6 \times 10^6$	cm s^{-2}
x_0	$= 0.04$	cm
h	$= 0.1$	cm
τ	$= 1 \times 10^{-4}$	s
M	$= 5 \times 10^{-3}$	g cm^{-2}
D	$= 5 \times 10^{-3}$	s cm^{-1}
α	$= 1.2 \times 10^6$	s^{-2}
β	$= 1.5 \times 10^3$	s^{-1}
f	$= 2500$	Hz
d_1	$= 0.1$	cm
q	$= 0.175$	-
L	$= 2.0$	cm
γ	$= 0.9$	-
c_a	$= 3.5 \times 10^4$	cm s^{-1}
c_h	$= 6.09 \times 10^4$	cm s^{-1}
ρ_a	$= 0.00114$	g cm^{-3}
ρ_h	$= 0.00034$	g cm^{-3}

Tabla A.1: Valores de parámetros estándar para el Capítulo 5, a menos que se indique lo contrario. Los valores de parámetros son iguales para las dos fuentes, a menos que se indique lo contrario.

Bibliografía

- [Abarbanel et al. 2003] H. D. I. Abarbanel, L. Gibb, G. B. Mindlin and S. Talathi. Mapping neural architectures onto acoustic features of birdsong. *J. Neurophysiol.* **92**, 96–110.
- [Amador et al. 2005] A. Amador, M. A. Trevisan and G. B. Mindlin. Simple neural substrate predicts complex rhythmic structure in duetting birds. (in preparation)
- [Arfken and Weber 1995] G. B. Arfken and H. J. Weber. *Mathematical methods for physicists*. Academic Press, San Diego.
- [Ballintijn and ten Cate 1998] M. R. Ballintijn and C. ten Cate. Sound production in the collared dove: a test of the whistle hypothesis. *J. Exp. Biol.* **201**, 1637–1649.
- [Beckers et al. 2003a] G. J. L. Beckers, R. A. Suthers, and C. ten Cate. Pure-tone birdsong by resonance filtering of harmonic overtones. *Proc. Natl. Acad. Sci. USA* **100**, 7372–7376.
- [Beckers et al. 2003b] G. J. L. Beckers, R. A. Suthers, and C. ten Cate. Mechanisms of frequency and amplitude modulation in the ring dove song. *J. Exp. Biol.* **206**, 1833–1843.
- [Beil 1962] R. Beil. Frequency analysis of vowels produced in a helium-rich environment. *J. Acoust. Soc. Am.* **34**, 347–349.
- [Berry and Titze 1996] D. A. Berry and I. R. Titze. Normal modes in a continuum model of vocal fold tissues. *J. Acoust. Soc. Am.* **100**, 3345–3354.
- [Bordelois 2003] I. Bordelois. *La palabra amenazada*. Libros del Zorzal, Buenos Aires. ISBN 987-10-8116-2.
- [Bottjer 2002] S. W. Bottjer. Neural strategies for learning during sensitive periods of development. *J. Comp. Physiol. A* **188**, 917–928.
- [Brackenbury 1979] J. H. Brackenbury. Aeroacoustics of the vocal organ in birds. *J. Theor. Biol.* **81**, 341–349.
- [Brainard and Doupe 2002] M. S. Brainard and A. J. Doupe. What songbirds teach us about learning. *Nature* **417**, 351–358.
- [Brenowitz et al. 1997] E. A. Brenowitz, D. Margoliash and K. W. Nordeen. An introduction to birdsong and the avian song system. *J. Neurobiol.* **33**, 495–500.

- [Brittan-Powell et al. 1997] E. F. Brittan-Powell, R. J. Dooling, O. N. Larsen, and J. T. Heaton. Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*). *J. Acoust. Soc. Am.* **101**, 578–589.
- [Burger 1980] J. Burger. El canto a dúo del hornero (*Furnarius rufus*). *El Hornero*, **12**, 17–21.
- [Calder 1970] W. A. Calder. Respiration during song in the canary (*Serinus canaria*). *Comp. Biochem. Physiol.* **32**, 251–258.
- [Casey and Gaunt 1985] R. M. Casey and A. S. Gaunt. Theoretical models of the avian syrinx. *J. Theor. Biol.* **116**, 45–64.
- [Doupe and Kuhl 1999] A. J. Doupe and P. K. Kuhl. Birdsong and human speech: Common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631.
- [DRAE] Diccionario de la Real Academia Española, 22a. edición (2001). <http://www.rae.es/>
- [Drăgănoiu et al. 2002] T. I. Drăgănoiu, L. Nagle and M. Kreutzer. Directional female preference *Proc. R. Soc. Lond. B* **269**, 2525–2531.
- [Elemans et al. 2003] C. P. H. Elemans, O. N. Larsen, M. R. Hoffmann and J. L. van Leeuwen. Quantitative modeling of the biomechanics of the avian syrinx. *Anim. Biol.* **53**, 183–193.
- [Elemans 2004] C. P. H. Elemans. *How do birds sing? Sound analysis - Mechanical modelling - Muscular control*. PhD thesis, Wageningen University, Wageningen.
- [Elemans et al. 2004] C. P. H. Elemans, I. L. Y. Spierts, U. K. Müller, J. L. Leeuwen and F. Goller. Superfast muscles control dove's trill. *Nature* **431**, 146.
- [Fant 1960] G. Fant. *Acoustic Theory of Speech Production*. Mouton, The Hague, The Netherlands.
- [Farries 2001] M. A. Farries. The oscine song system considered in the context of the avian brain: lessons learned from comparative neurobiology. *Brain Behav. Evol.* **58**, 80–100.
- [Farries 2004] M. A. Farries. The avian song system in comparative perspective. *Ann. N. Y. Acad. Sci.* **1016**, 61–76.
- [Fee et al. 1998] M. S. Fee, B. Shraiman, B. Pesaran and P. P. Mitra. The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature* **395**, 67–71.
- [Fee 2002] M. S. Fee. Measurement of the linear and nonlinear mechanical properties of the oscine syrinx: implications for function. *J. Comp. Physiol. A* **188**, 829–839.
- [Flanagan and Landgraf 1968] J. L. Flanagan and L. Landgraf. Self-oscillating source for vocal-tract synthesizers. *IEEE Trans. Audio Electroacoust.* **AU-16**, 57.

- [Fletcher 1988] N. H. Fletcher, Bird song—a quantitative acoustic model. *J. Theor. Biol.* **135**, 455–481.
- [Fletcher 2000] N. H. Fletcher. A class of chaotic bird calls? *J. Acoust. Soc. Am.* **108**, 821–826.
- [Fletcher and Tarnopolsky 1999] N. H. Fletcher and A. Tarnopolsky. Acoustics of the avian vocal tract. *J. Acoust. Soc. Am.* **105**, 35–49.
- [Fry 1998] C. L. Fry. *A source-filter model of birdsong production*. Ph.D. Thesis, University of California, San Diego, USA.
- [Gardner et al. 2001] T. Gardner, G. Cecchi, M. Magnasco, R. Laje and G. B. Mindlin. Simple motor gestures for birdsongs. *Phys. Rev. Lett.* **87**, 208101–1–4.
- [Gaunt et al. 1982] A. S. Gaunt, S. L. L. Gaunt, and R. M. Casey, Syringeal mechanisms reassessed: evidence from *Streptopelia*. *Auk* **99**, 474–494.
- [Gaunt 1983] A. S. Gaunt. An hypothesis concerning the relationship of syringeal structure to vocal abilities. *Auk* **100**, 853–862.
- [Gaunt and Gaunt 1985] A. S. Gaunt and S. L. L. Gaunt. Electromyographics studies of the syrinx in parrots (Aves, Psittacidae). *Zoomorph.* **105**, 1–11.
- [Glass 2001] L. Glass. Synchronization and rhythmic processes in physiology. *Nature* **410**, 277–284.
- [Goller and Cooper 2004] F. Goller and B. G. Cooper. Peripheral motor dynamics of song production in the zebra finch. *Ann. N. Y. Acad. Sci.* **1016**, 130–152.
- [Goller and Daley 2001] F. Goller and M. A. Daley. Novel motor gestures for phonation during inspiration enhance the acoustic complexity of birdsong. *Proc. R. Soc. Lond. B* **268**, 2301–2305.
- [Goller and Larsen 1997a] F. Goller and O. N. Larsen. In situ biomechanics of the syrinx and sound generation in pigeons. *J. Exp. Biol.* **200**, 2165–2176.
- [Goller and Larsen 1997b] F. Goller and O. N. Larsen. A new mechanism of sound generation in songbirds. *Proc. Natl. Acad. Sci. USA* **94**, 14787–14791.
- [Goller and Suthers 1995] F. Goller and R. A. Suthers. Implications for lateralization of bird song from unilateral gating of bilateral motor patterns. *Nature* **373**, 63–66.
- [Goller and Suthers 1996a] F. Goller and R. A. Suthers. Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J. Neurophysiol.* **75**, 867–876.
- [Goller and Suthers 1996b] F. Goller and R. A. Suthers. Role of syringeal muscles in controlling the phonology of bird song. *J. Neurophysiol.* **76**, 287–300.
- [Goller and Suthers 1999] F. Goller and R. A. Suthers. Bilaterally symmetrical respiratory activity during lateralized birdsong. *J. Neurobiol.* **41**, 513–523.

- [González and Piro 1983] D. L. González and O. Piro. Chaos in a nonlinear driven oscillator with exact solution. *Phys. Rev. Lett.* **50**, 870–872.
- [Gray 1963] Gray D. E. (editor), *American Institute of Physics Handbook* (2nd edition). McGraw-Hill, New York.
- [Greenewalt 1968] C. H. Greenewalt. *Bird song: acoustics and physiology*. Smithsonian Institution Press, Washington DC.
- [Guckenheimer and Holmes 1996] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields* (5th printing). Springer-Verlag, New York.
- [Hahnloser et al. 2002] R. H. R. Hahnloser, A. A. Kozhevnikov, and M. S. Fee. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**, 65–70.
- [Hale 1977] J. K. Hale. *Theory of functional differential equations*. Springer, New York.
- [Hartley and Suthers 1989] R. S. Hartley and R. A. Suthers. Airflow and pressure during canary song: direct evidence for mini-breaths. *J. Comp. Physiol. A* **165**, 15–26.
- [Hegger et al. 1998] R. Hegger, M. J. Bünner, H. Kantz and A. Giaquinta. Identifying and modeling delay feedback systems. *Phys. Rev. Lett.* **81**, 558–561.
- [Hérissant 1753] F. D. Hérissant. Recherches sur les organes de la voix des quadrupèdes et de celle des oiseaux. *Mém. l'Acad. Sci. Paris*, 279–295.
- [Hoese et al. 2000] W. J. Hoese, J. Podos, N. C. Boetticher and S. Nowicki. Vocal tract function in birdsong production: experimental manipulation of beak movements. *J. Exp. Biol.* **203**, 1845–1855.
- [Hoppensteadt and Izhikevich 1997] F. C. Hoppensteadt and E. M. Izhikevich. *Weakly connected neural networks*. Springer-Verlag, New York.
- [Hunter 1980] M. L. Hunter. Vocalization during inhalation in a nightjar. *Condor* **82**, 101–103.
- [Ikeda and Matsumoto 1987] K. Ikeda and K. Matsumoto. High-dimensional chaotic behavior in systems with time-delayed feedback. *Physica D* **29**, 223–235.
- [Ishizaka and Flanagan 1972] K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Sys. Tech. Journal* **51**, 1233–1268.
- [Jarvis et al. 2000] E. D. Jarvis, S. Ribeiro, M. L. da Silva, D. Ventura, J. Vieillard and C. V. Mello. Behaviourally driven gene expression reveals song nuclei in hummingbird brain. *Nature*, **406**, 628–632.
- [Jarvis et al. 2002] E. D. Jarvis, V. A. Smith, K. Wada, M. V. Rivas, M. McElroy, T. V. Smulders, P. Carninci, Y. Hayashizaki, F. Dietrich, X. Wu, P. McConnell, J. Yu, P. P. Wang, A. J. Hartemink and S. Lin. A framework for integrating the songbird brain. *J. Comp. Physiol. A* **188**, 961–980.

- [Jarvis 2004] E. D. Jarvis. Learned birdsong and the neurobiology of human language. *Ann. N. Y. Acad. Sci.* **1016**, 749–777.
- [Keener and Sneyd 1998] J. Keener and J. Sneyd. *Mathematical physiology*. Springer, New York.
- [Kinsler et al. 1982] L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders. *Fundamentals of Acoustics*. John Wiley & Sons, New York.
- [Klatt and Stefanski 1974] D. H. Klatt and R. A. Stefanski. How does a mynah bird imitate human speech? *J. Acoust. Soc. Am.* **55**, 822–832.
- [Laje et al. 2001] R. Laje, T. Gardner and G. B. Mindlin. Continuous model for vocal fold oscillations to study the effect of feedback. *Phys. Rev. E* **64**, 056201-1–7.
- [Laje et al. 2002] R. Laje, T. J. Gardner and G. B. Mindlin. Neuromuscular control of vocalizations in birdsong: a model. *Phys. Rev. E* **65**, 051921-1–8.
- [Laje and Mindlin 2002] R. Laje and G. B. Mindlin. Diversity within a birdsong. *Phys. Rev. Lett.* **89**, 288102-1–4.
- [Laje and Mindlin 2003] R. Laje and G. B. Mindlin. Highly structured duets in the song of the south american hornero. *Phys. Rev. Lett.* **91**, 258104-1–4.
- [Laje and Mindlin 2005] R. Laje and G. B. Mindlin. Modeling source-source and source-filter acoustic interaction in birdsong. *Phys. Rev. E* (in press).
- [Larsen and Goller 1999] O. N. Larsen and F. Goller. Role of syringeal vibrations in bird vocalizations. *Proc. R. Soc. Lond. B* **266**, 1609–1615.
- [Larsen and Goller 2002] O. N. Larsen and F. Goller. Direct observation of syringeal muscle function in songbirds and a parrot. *J. Exp. Biol.* **205**, 25–35.
- [Lucero 1998] J. C. Lucero. A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset offset. *J. Acoust. Soc. Am.* **105**, 423–431.
- [Marler 1970] P. A. Marler. A comparative approach to vocal learning: song development in white-crowned sparrows. *J. Comp. Physiol. Psychol.* **71**, 1–25.
- [McCasland 1987] J. S. McCasland. Neuronal control of bird song production. *J. Neurosci.* **7**, 23–39.
- [Mindlin et al. 2003] G. B. Mindlin, T. J. Gardner, F. Goller and R. Suthers. Experimental support for a model of birdsong production. *Phys. Rev. E* **68**, 041908.
- [Mindlin and Laje 2005] G. B. Mindlin and R. Laje. *The Physics of Birdsong*. Springer, New York. ISBN 3-540-25399-8.
- [Miskimen 1951] M. Miskimen. Sound production in passerine birds. *Auk* **68**, 493–504.

- [Nottebohm 1969] F. Nottebohm. The song of the chingolo *Zonotrichia capensis* in Argentina: description and evaluation of a system of dialects. *Condor* **71**, 299–315.
- [Nottebohm et al. 1976] F. Nottebohm, T. M. Stokes, C. M. Leonard. Central control of song in the canary. *J. Comp. Neurol.* **165**, 457–486.
- [Nowicki and Capranica 1986a] S. Nowicki and R. R. Capranica. Bilateral syringeal interaction in the production of an oscine bird sound. *Science* *231*, 1297–1299.
- [Nowicki and Capranica 1986b] S. Nowicki and R. R. Capranica. Bilateral syringeal coupling during phonation of a songbird. *J. Neurosci.* *6*, 3595–3610.
- [Nowicki 1987] S. Nowicki. Vocal tract resonances in oscine bird sound production: Evidence from birdsongs in a helium atmosphere. *Nature* **325**, 53–55.
- [Nowicki and Searcy 2004] S. Nowicki and W. A. Searcy. Song function and the evolution of female preferences —Why birds sing, why brains matter. *Ann. N. Y. Acad. Sci.* **1016**, 704–723.
- [Pelorson et al. 1994] X. Pelorson, A. Hirschberg, R. R. van Hassel, A. P. J. Wijnands and Y. Auregan. Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model. *J. Acoust. Soc. Am.* **96**, 3416–3431.
- [Perkel 2004] D. J. Perkel. Origin of the Anterior Forebrain Pathway. *Ann. N. Y. Acad. Sci.* **1016**, 736–748.
- [Podos 1996] J. Podos. Motor constraints on vocal development in a songbird. *Anim. Behav.* **51**, 1061–1070.
- [Podos 1997] J. Podos. A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution* **51**, 537–551.
- [Poole et al. 2005] J. H. Poole, P. L. Tyack, A. S. Stoeger-Horwath, S. Watwood. Animal behaviour: Elephants are capable of vocal learning. *Nature* **434**, 455–456.
- [Rothenberg 1986] M. Rothenberg. Così fan tutte, and what it means, in *Vocal fold physiology: Laryngeal function in phonation and respiration*. T. Baer, C. Sasaki and K. S. Harris, eds. College Hill Press, San Diego.
- [Sciamarella et al. 2005] D. Sciamarella, J. Zanella, R. Laje and G. B. Mindlin. Bilateral source interaction in a syrinx model of an oscine bird. (en preparación)
- [Slabbekoorn and Peet 2002] H. Slabbekoorn and M. Peet. Birds sing at a higher pitch in urban noise. *Nature* **424**, 267.
- [Spiro et al. 1999] J. E. Spiro, M. B. Dalva and R. Mooney. Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *J. Neurophysiol.* **81**, 3007–3020.

- [Steinecke and Herzel 1995] I. Steinecke and H. Herzel. Bifurcations in an asymmetric vocal-fold model. *J. Acoust. Soc. Am.* **97**, 1874–1884.
- [Story and Titze 1995] B. H. Story and I. R. Titze. Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Am.* **97**, 1249–1260.
- [Straneck 1990a] R. Straneck. *Canto de las Aves Pampeanas I*. LOLA, Buenos Aires.
- [Straneck 1990b] R. Straneck. *Canto de las Aves Pampeanas II*. LOLA, Buenos Aires.
- [Straneck 1990c] R. Straneck. *Canto de las Aves de Misiones I*. LOLA, Buenos Aires.
- [Strogatz 1994] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. Perseus, Cambridge.
- [Suthers 1994] R. A. Suthers. Variable asymmetry and resonance in the avian vocal tract: a structural basis for individually distinct vocalizations. *J. Comp. Physiol. A* **75**, 457–466.
- [Suthers 1997] R. A. Suthers. Peripheral control and lateralization of birdsong. *J. Neurobiol.* **33**, 632–652.
- [Suthers 2001] R. A. Suthers. Peripheral vocal mechanisms in birds: are songbirds special? *Neth. J. Zool.* **51**, 217–242.
- [Suthers et al. 1999] R. A. Suthers, F. Goller and C. Pytte. The neuromuscular control of birdsong. *Phil. Trans. R. Soc. Lond. B* **354**, 927–939.
- [Suthers and Margoliash 2002] R. A. Suthers and D. Margoliash. Motor control of birdsong. *Curr. Opin. Neurobiol.* **12**, 684–690.
- [Suthers and Zollinger 2004] R. A. Suthers and S. A. Zollinger. Producing song: The vocal apparatus. *Ann. N. Y. Acad. Sci.* **1016**, 109–129.
- [Švec et al. 1996] J. G. Švec, H. K. Schutte and D. G. Miller. A subharmonic vibratory pattern in normal vocal folds. *J. Speech Hear. Res.* **39**, 135–143.
- [Thorpe 1961] W. H. Thorpe. *Bird-Song*. Cambridge University Press, Cambridge.
- [Titze 1973] I. R. Titze. The human vocal cords: a mathematical model (part I). *Phonetica* **28**, 129–170.
- [Titze 1974] I. R. Titze. The human vocal cords: a mathematical model (part II). *Phonetica* **29**, 1–21.
- [Titze 1984] I. R. Titze. Parametrization of the glottal area, glottal flow, and vocal fold contact area. *J. Acoust. Soc. Am.* **75**, 570–580.
- [Titze 1988] I. R. Titze. The physics of small-amplitude oscillation of the vocal folds. *J. Acoust. Soc. Am.* **83**, 1536–1551.

- [Titze and Story 1997] I. R. Titze and B. H. Story. Acoustic interactions of the voice source with the lower vocal tract. *J. Acoust. Soc. Am.* **101**, 2234–2243.
- [Tubaro et al. 1993] P. L. Tubaro, E. T. Segura and P. Handford. Geographic variation in the song of the rufous-collared sparrow in eastern Argentina. *Condor* **95**, 588–595.
- [Vicario 1991] D. S. Vicario. Contributions of syringeal muscles to respiration and vocalization in the zebra finch. *J. Neurobiol.* **22**, 63–73.
- [Vu et al. 1994] E. T. Vu, M. E. Mazurek and Y. C. Kuo. Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neurosci.* **14**, 6924–6934.
- [Wild 1997] J. M. Wild. Neural pathways for the control of birdsong production. *J. Neurobiol.* **33**, 653–670.
- [Wilden et al. 1998] I. Wilden, H. Herzel, G. Peters and G. Tembrock. Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics* **9**, 171–196.
- [Yu and Margoliash 1996] A. C. Yu and D. Margoliash. Temporal hierarchical control of singing in birds. *Science* **273**, 1871–1875.
- [Zollinger and Suthers 2004] S. A. Zollinger and R. A. Suthers. Motor mechanisms of a vocal mimic: Implications for birdsong production. *Proc. R. Soc. London B* **271**, 483–491.

Agradecimientos

Las palabras de repente huyen cada vez que vuelvo a esta página en blanco. El blanco se llena con un álbum de fotos tan veloz y apretado que no me da tiempo a escribir. Algunas de las personas en ese álbum figuran aquí.

En primer lugar quiero agradecer a mi director, Gabriel Mindlin. En cada charla, en cada sugerencia, en cada párrafo me sentí inmerso en la profunda atmósfera científica que Gabriel crea a su alrededor. Con una pizca que pueda llevarme de todo eso, ya me alcanza. Y eso no es todo. Su desinteresada dedicación a nosotros, sus estudiantes, es admirable: es un gran director, y una gran persona. Gabo, gracias por todo, y muy especialmente gracias por la oportunidad.

Ale Yacomotti, compañero de estudios y, apenas después, compañero de la vida.

Manuel Eguía, que me enseña muchas cosas aunque él no lo sospeche.

Marcos Trevisan y Pablo Jercog, compañeros queridos y queridísimos amigos. A uno ya lo extraño, al otro ya lo voy a extrañar.

Los chicos del Laboratorio de Sistemas Dinámicos: Jorge Méndez, Dani Zysman, Ana Amador, Jacobo Sitt, Mariano Gabitto, Guadalupe García. Mucho brindis en este grupo, eh. Da gusto trabajar con ustedes.

Tim J. Gardner, con quien compartí los primeros pasos en el proyecto de las aves.

Muchas personas a quienes tuve el placer de conocer, y que de forma directa o indirecta ahora forman parte de mi doctorado: entre ellos Juan Pablo Paz, Guillermo Dussel, Guillermo Cecchi, Silvia Loza Montaña. De todos me llevo algo para atesorar.

Mi familia: Jimena, Santiago, Pablo y Kanky, por quién sentirme más orgulloso. Son un regalo del cielo, como Laura.

Siempre lo digo... no, Marcos? Cuando sea grande, quiero ser primero como Manuel y después como Gabriel.