



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

**Programa interactivo para analizar la voz cantada
mediante técnicas de procesamiento digital de señales**

T E S I S

QUE PARA OBTENER EL GRADO DE
MAESTRÍA EN MÚSICA
CAMPO DE CONOCIMIENTO: TECNOLOGÍA MUSICAL

PRESENTA:

ING. GISELA GRACIDA OLVERA

TUTOR: DR. FELIPE ORDUÑA BUSTAMANTE

México, D. F, octubre de 2010.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

CONTENIDO

	Página
Dedicatoria - - - - -	3
Agradecimientos - - - - -	4
Introducción - - - - -	5
Capítulo 1. ANTECEDENTES - - - - -	6
1.1 La enseñanza de la técnica del canto - - - - -	6
1.2 Investigaciones previas acerca de la voz cantada - - - - -	12
1.3 Modelo físico y análisis acústico de la señal de voz - - - - -	16
1.4 Parámetros acústicos característicos de la voz cantada - - - - -	21
1.4.1. Frecuencia fundamental - - - - -	21
1.4.2. Formante del cantante - - - - -	23
1.4.3. Sintonía formante – armónico - - - - -	25
1.4.4. Vibrato - - - - -	25
1.5 Programas existentes para el análisis de audio y voz - - - - -	26
Capítulo 2. ELABORACIÓN DEL PROGRAMA INTERACTIVO “EVOCANTO” - - - - -	29
2.1 Análisis de requerimientos - - - - -	29
2.2 Herramientas de programación - - - - -	30
2.2.1 Python - - - - -	30
2.2.2 pySndObj - - - - -	32
2.3 Diseño y desarrollo del programa - - - - -	33
2.3.1 Conformación de la señal de entrada - - - - -	33
2.3.2 Análisis espectral de la señal de voz cantada - - - - -	35
2.3.3 Presentación de los resultados - - - - -	39
Capítulo 3. USO Y EVALUACIÓN DEL PROGRAMA “EVOCANTO” - - - - -	42
3.1 Guías de uso y sugerencias didácticas - - - - -	42
3.2 Pruebas técnicas y ejemplos de aplicación - - - - -	44
3.3 Pruebas de opinión - - - - -	50
Conclusiones - - - - -	53
Referencias - - - - -	55
Apéndice A. Programas existentes para el análisis de audio y voz - - - - -	58
Apéndice B. Código fuente del programa Evocanto - - - - -	63
Apéndice C. Ventana de Ayuda del programa Evocanto - - - - -	68

A mi hija,

Significas para mi grandes cosas.
Amas en forma total y sincera,
Razonas siempre a tu manera,
Alivia y angustias mis horas.

Risa sembrada al amanecer.
Unes el mundo con verdad,
Bonanza, canto y tempestad.
Iluminas la vida al crecer.

***** GRACIAS *****

- por el apoyo, la ayuda, el refugio, la comprensión, el aliento y la motivación:
 - a la Universidad Nacional Autónoma de México, a la Escuela Nacional de Música y al Centro de Ciencias Aplicadas y Desarrollo Tecnológico; a las personas que con su labor diaria conforman estas instituciones.
 - a mi familia: Ma. Del Pilar, Gabriel, Gabo, Isela y Sara Rubí.
 - a Rosa Evelia.
 - muy especialmente al Dr. Felipe Orduña Bustamante; nuestra Tesis.
 - al Padre Eterno, a la vida, a la música y a la ciencia.

- por la beca otorgada por la DGEP-UNAM para cursar la Maestría
- por el financiamiento recibido por la DGAPA-UNAM, al proyecto PAPIME PE-401207
- por el financiamiento de “Fomento a la Graduación” otorgado por la DGEP-UNAM

Gisela Gracida Olvera

INTRODUCCIÓN

Los avances en la investigación científica acerca de la voz, han permitido en las últimas tres décadas el desarrollo de nuevos procedimientos para evaluar la calidad de la voz cantada. En paralelo a ello, surge la necesidad de proporcionar instrumentos auxiliares del análisis, que puedan aprovechar las capacidades generalizadas de las computadoras personales y portátiles, sin tener que adquirir aparatos más especializados y, así, sean cada vez más accesibles a los cantantes.

Esta tesis trata acerca de la creación y desarrollo de un programa de cómputo amigable, para analizar la voz cantada, construido con base en las técnicas de procesamiento digital de señales. Se presenta como una herramienta tecnológica de contribución a la evaluación y mejoramiento de la técnica vocal de los estudiantes de canto; que pueda ser utilizado como un auxiliar del profesor, con la intención de promover en los alumnos la asimilación significativa de elementos en su técnica vocal, fomentar el logro de objetivos y permitir el desarrollo de estrategias de exploración en su aprendizaje.

La organización de este trabajo consta de cuatro capítulos.

El Capítulo 1 trata los conceptos que están relacionados con la enseñanza de la técnica del canto; describe las investigaciones anteriores de las cuales se desprende el desarrollo de la presente; también contiene una breve descripción del modelo físico y el análisis espectral de la producción vocal, así como de los parámetros de la voz cantada que serán evaluados en este trabajo de tesis; y, por último, enumera los tipos de programas de cómputo que existen en la actualidad para el análisis de señales de voz y para entrenamiento en el canto.

En el Capítulo 2 se realiza el análisis de los requerimientos del programa denominado *Evocanto*, desarrollado en esta tesis. Además, se justifica la selección de las herramientas de programación con las cuales se elaboró. Posteriormente se describe detalladamente el diseño y funcionamiento del programa.

En el Capítulo 3 se muestran algunas pruebas del programa *Evocanto* en su implementación real; se proponen algunas aplicaciones didácticas; también se presentan los resultados de la evaluación del funcionamiento del programa *Evocanto*, a partir de su utilización por parte de estudiantes y de profesores de canto; y mediante la elaboración y aplicación de cuestionarios de opinión.

Posteriormente, se describen a manera de conclusiones, los alcances y aportaciones de este trabajo, así como sus posibles ampliaciones futuras.

CAPÍTULO 1. ANTECEDENTES

El sustento teórico necesario para desarrollar y comprender los capítulos posteriores, se presenta en este capítulo. Se integran los siguientes elementos: la técnica vocal en la enseñanza del canto; algunos temas de estudio y análisis de la voz cantada; la acústica de la voz cantada, los parámetros correspondientes de análisis acústico; y los programas de cómputo que ya existen, diseñados para hacer análisis digital de audio, de voz y para entrenamiento de la voz cantada.

1.1 La enseñanza de la técnica del canto

El canto es un arte que implica diversas áreas en su desarrollo y ejecución, desde la anatomía y la fisiología hasta la psicoacústica y la estética. En el contexto de la tradición académica occidental, la formación de un cantante profesional requiere muchos años de estudio y disciplina, durante los cuales debe aprender a dominar el arte de la interpretación musical. Esto conlleva el desarrollo de habilidades para conjuntar en inteligencia y emoción, las características de distintos estilos y formas musicales; a través del estudio de la técnica vocal, así como de los necesarios aspectos musicales, históricos, lingüísticos e histriónicos.

La técnica vocal corresponde a la parte de preparación física del cantante; a la coordinación del funcionamiento natural de la voz, por medio de la cual se adquiere la habilidad de cantar cómodamente en amplios intervalos, se corrigen defectos vocales y se homogeniza la voz en todos sus registros¹. El entrenamiento también puede permitir al cantante producir un sonido acústicamente más eficiente, en términos de utilizar el consumo mínimo de energía para que la intensidad sonora resultante sea mayor y, así, llenar un espacio físico grande o sobresalir del acompañamiento de los instrumentos musicales, sin la necesidad de una amplificación artificial.

¹ Según Johan Sundberg, la definición más común de registro es: "un rango de frecuencias fonadas en el cual todos los tonos son percibidos como producidos de una forma semejante y con un timbre semejante". Varios autores como Van den Berg (1968), Large (1973) y Hollien (1971) describen la existencia de tres registros básicos o principales: el **glottal fry** (registro de frote glotal o registro de pulso); el registro **modal** que incluye los denominados "registro medio" y "de cabeza"; el **falsete** y un cuarto registro, el registro **flauta** o silbido, en la zona de agudos.

Cada persona posee una voz especial, un timbre individual, en esencia determinado por la constitución y características físicas del propio aparato vocal. Este hecho es de suma importancia en el desarrollo y el estudio de la técnica, pues otorga a cada alumno posibilidades muy particulares. Es por ello que, aunque los principios y fundamentos de la técnica vocal sean los mismos en general, el profesor debe adaptar su enseñanza a cada alumno en particular.

El conocimiento del aparato vocal y, sobre todo, su funcionamiento para la producción de voz, es un tópico básico e indispensable para comenzar el estudio de la técnica vocal. Es equivalente a la noción que debe tener un instrumentista acerca de su instrumento musical.

El aparato vocal está constituido por tres sistemas principales, que son: el sistema respiratorio, el sistema de fonación y el sistema de resonancia. En ese mismo orden es como se produce la voz. En el sistema respiratorio se genera el flujo de aire, que pasa por el sistema de fonación, donde se produce el sonido y, por último, le son dadas cualidades de timbre, calidad vocal y amplitud, a través del sistema de resonancia (figura 1.1).

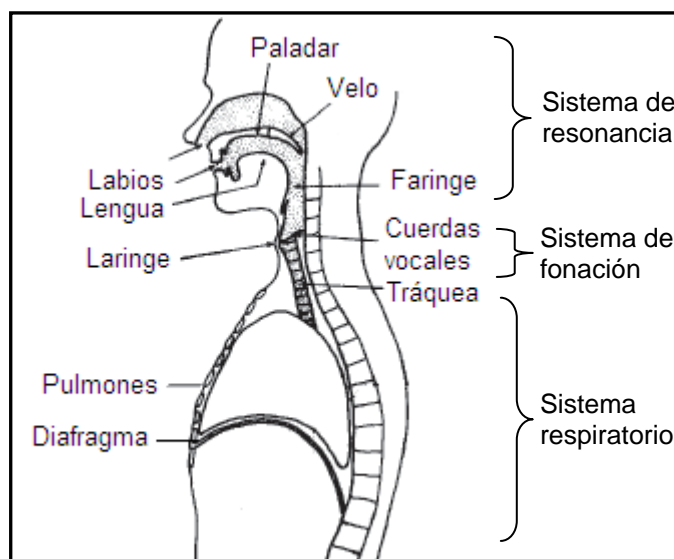


Figura 1.1 Esquema del aparato vocal

Al sistema respiratorio lo conforman principalmente los pulmones y el músculo diafragma, que son los responsables de almacenar el aire y controlar la respiración, respectivamente. El movimiento del diafragma obliga a los pulmones a ampliarse y contraerse, causando los cambios de presión necesarios para la inspiración y la

espiración. La liberación de la presión almacenada en los pulmones proporciona la corriente de aire necesaria para la producción del sonido.

La laringe constituye el principal órgano de fonación. Es un esqueleto de cartílagos (tiroideos, cricoides, y aritenoides) que incluye y apoya a unas estructuras de músculo y ligamentos cubiertos por membranas mucosas. Estas estructuras son los pliegues o cuerdas vocales, que son la fuente primaria para la producción de la voz.

Cuando los pliegues son separados o abducidos, permiten al aire pasar libremente entre ellos, como en el caso de la respiración. Cuando los pliegues son reunidos, o aducidos, la corriente de aire es oprimida, que es la condición preparatoria para la vibración. Los músculos de los pliegues vocales pueden cambiar la forma y la rigidez de los pliegues, causando cambios en la producción del sonido.

El sistema de resonancia es la combinación de cavidades por encima de la laringe que comprende la faringe, la cavidad bucal y la cavidad nasal². Al área comprendida entre la parte superior de la laringe y los labios, se le llama tracto vocal.

El tracto vocal tiene la capacidad de asumir una amplia gama de formas diferentes, que fácilmente son cambiadas por la articulación (modificación de la posición del velo del paladar, la mandíbula, la lengua y los labios), lo que proporciona una gran flexibilidad de afectación al sonido resultante.

Diversas partes del sistema vocal, como las propias cuerdas vocales, la glotis y el diafragma, no son accesibles a través de sensaciones nerviosas, es por ello que el profesor de canto utiliza métodos indirectos que influyen la voz de su alumno, como el utilizar la imaginación, ser consciente de sensaciones musculares en órganos adyacentes y el desarrollo de la atención a su propio sonido producido. Estas técnicas han sido diseñadas para cultivar la conciencia y un control preciso sobre las estructuras vocales para obtener una óptima resonancia, potencia y vibrato [Sundberg, 2000].

Los principios de la técnica vocal, que los profesores de canto promueven en sus alumnos, son los que se describen a continuación.

² Vennard describe que en 1954, Wooldridge intentó determinar la contribución de las cavidades nasales a la voz cantada, al comparar las vocales producidas por seis cantantes profesionales en dos condiciones: normal y con las cavidades nasales llenas de gasa de algodón. Él no encontró diferencias significativas entre los espectros de las vocales producidas en las dos condiciones, y un jurado de oyentes expertos fue incapaz de distinguir las dos condiciones al escuchar las grabaciones en cinta. Wooldridge concluyó que el término "resonancia nasal" no tiene validez en la descripción de la calidad de la voz cantada. Una repetición de este experimento por cinco cantantes masculinos, incluyendo al propio Vennard, confirmó las conclusiones originales [Vennard, 1967].

La postura corporal.

Es esencial para el cantante adquirir una buena postura corporal, ya que está directamente relacionada con la emisión de la voz. Es común que en las clases de técnica se tenga a mano un espejo donde el alumno podrá observar los cambios en su postura del cuerpo.

La postura correcta habla por sí misma del cantante. Ya que los cantantes pueden mejorar bastante su entrada de aliento en una postura correcta, es casi inevitable que ellos canten mucho mejor también de esa manera [Roe, 1970].

La enseñanza tradicional del canto menciona una postura “noble” que significa erguida, bien apoyada en el suelo, pero sin tensiones innecesarias. Para comenzar a practicar, se recomienda poner especial atención en las siguientes partes del cuerpo [Collins, 1999]:

Pies. Distribuir el peso del cuerpo equitativamente en los talones de los pies, que deben estar bien apoyados en el piso y tener una separación entre sí, de la misma proporción que el ancho de las caderas.

Rodillas. Jamás deben estar juntas.

Abdomen. La parte alta del abdomen, donde se encuentra el diafragma, por debajo de las costillas, donde termina el esternón, debe ser un área que permanezca flexible y responder suavemente al ciclo de la respiración.

Tórax. Debe estar erguido, pero no tenso.

Hombros. Hacia atrás, hacia abajo y relajados.

Cuello. No se debe acumular ninguna tensión en todo el perímetro del cuello.

Brazos y manos. Relajados, a los lados del cuerpo.

Otro factor que es exigente de la postura corporal, es la interacción que existe entre el cantante y su público. Actualmente, se han incorporado conocimientos de distintas técnicas específicas, como la técnica Alexander [Heirich, 2005] o el método Feldenkrais [Blades-Zeller, 2002] que tratan más de la correcta utilización del cuerpo en movimiento que de una postura en concreto.

La facilidad, potencia, ductilidad y calidad sonora de la voz, son producidas por determinados y específicos movimientos musculares de los órganos implicados en la emisión, por lo que el profesor de canto muestra al alumno una técnica gestual. Una buena capacidad de introspección y experiencia escénica le serán útiles para pormenorizar al alumno las sensaciones que va a experimentar [Ferrer, 2008].

La respiración.

Es la base fundamental de la correcta emisión de voz.

Existen dos factores que deben ser considerados en la respiración de los cantantes: la dirección y el control del aire [Roe, 1970].

La columna de aire es elevada y apoyada por medio de la presión hiperbárica pulmonar en dirección hacia el cráneo [Ferrer, 2008].

Durante la entrada de aire, los músculos abdominales están relajados y el pecho se expande; cuando el cantante va consumiendo su suministro de aire, debe evitar la contracción de las costillas, manteniéndolas elevadas, es decir, el cambio del contorno de la caja torácica no debe ser sensible a ello si la respiración es hecha correctamente.

La inspiración nasal debe ser profunda, suave y silenciosa. En la espiración, el cantante tiene que saber distribuir el aire inspirado por medio de una dosificación precisa, ya que si se deja escapar demasiado aire, la presión espiratoria que actúa sobre las cuerdas vocales será grande y elevará la altura del sonido y, además, se fatigará a las cuerdas vocales por la acción brusca de la apertura de la glotis, llamada "golpe de glotis". Pero si la presión espiratoria es débil, disminuye la altura del sonido, afectando la entonación exacta. La presión espiratoria influye en la intensidad, la entonación, la duración y el vibrato del sonido emitido [Canuyt, 1958].

La emisión y colocación de la voz.

El ataque, es decir, el comienzo del sonido, como resultado de un conjunto de movimientos naturales de las cuerdas vocales (aducción, tensión, presión espiratoria y oscilaciones vibratorias), se controla a través de la correcta respiración y por la postura corporal adecuada, sin tensiones musculares innecesarias. [Cornut, 1985].

Se busca una emisión fisiológicamente natural, cómoda, sin esfuerzo, con voz rotunda y clara. Para ello, la epiglotis y las paredes del orificio superior de la laringe se apartan bien para librar la salida de las vibraciones vocales; el velo del paladar se eleva y cierra la faringe nasal; la lengua es propulsada hacia delante con la punta aplicada a los incisivos inferiores; la boca deberá hallarse moderadamente abierta; la voz tendrá que estar situada hacia delante, sobre el paladar duro y los dientes, y el cantante buscará el timbre vocal "a flor de labios" [Canuyt, 1958].

Debe existir una sensación muy precisa del punto de emisión y mantenerse constante esa colocación y estabilidad laríngea. Según Ferrer, la coordinación perfecta entre un hiperbarismo pulmonar, aumentado por la tensión diafragmática, y la ligera

oclusión-retención del aire que precede al inicio de la actividad de las cuerdas, que se percibe como un pequeño chasquido, es la correcta colocación de la voz (*messa di voce*). Si se mantiene la voz en ese punto, permite, con la práctica, la emisión del sonido sin alteraciones, desde un *pianissimo* hasta un *forte* y viceversa [Ferrer, 2008].

El paso a través de los registros de voz.

Cuando la voz se traslada de un registro a otro, las cuerdas vocales cambian de espesor, longitud y tensión; también cambia el lugar clave de resonancia. Por lo tanto, el cantante tiene que hacer un cambio en la emisión. Las notas musicales donde se percibe claramente ese cambio, se llaman *notas de paso*, que, para el cantante, exigen un estudio especial, ya que con el entrenamiento se vuelven menos perceptibles. Cuando el cantante, a través de la práctica, logra suavizar, emparejar o fusionar el paso a través de los registros, se dice que la voz es *homogénea*, pues aparentemente no se perciben cambios en la emisión vocal.

La preparación indispensable para lograr efectuar el pasaje de la voz, según Ferrer, es el acto del bostezo (abertura relajada de la mandíbula), ya que en él existe una sincronización entre el movimiento de descenso laríngeo y el de la elevación del paladar. [Ferrer, 2008].

Articulación y pronunciación.

La correcta dicción, integrada por estos dos aspectos, proporciona la estética de la palabra cantada hacia la interpretación. La articulación entonada del lenguaje y la resonancia, se condicionan y complementan. Además de la exigencia para el cantante de distinguir las diferencias fonéticas de cada idioma, es necesaria la claridad en la proyección de las palabras, por lo que éstas deberán articularse siempre adelante y los labios deben estar muy activos, lo que permite la comprensión –por parte del escucha- del texto cantado [Cornut, 1985].

Es importante que los sonidos emitidos por la laringe no sean modificados en su calidad por alteraciones debidas a las consonantes [Ferrer, 2008].

La postura de la mandíbula debe ser siempre relajada, tanto en vocales como en consonantes [Roe, 1970].

Cuando los cantantes abren la boca de manera normal, los movimientos de mandíbula deben ser hacia abajo y atrás, suavemente, como si la cavidad sinovial y el cóndilo hubiesen sido aceitados (figura 1.1.2). La mandíbula funciona mecánicamente de

una manera maravillosamente equilibrada, y cuando el cantante permite el movimiento en sólo dos direcciones (abajo y arriba), la acción de la mandíbula será libre y sin fricción. De otra manera, si la mandíbula es empujada hacia fuera o a un lado, los músculos implicados en su operación se tensarán y el mecanismo de la articulación no funcionará correctamente. Es esencial, por lo tanto, que sea permitido a la mandíbula balancearse con facilidad extrema [Roe, 1970].

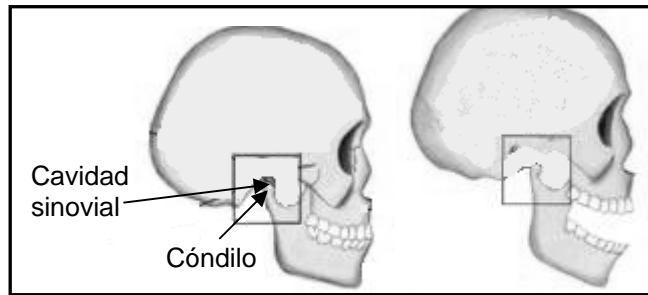


Figura 1.1.2 Movimiento libre de la mandíbula.

1.2 Investigaciones previas acerca de la voz cantada

La voz humana, como elemento en la música, ha sido objeto de inagotables investigaciones de tipo científico, que principalmente tratan los mecanismos fisiológicos que dan origen a la emisión vocal y las repercusiones de éstos en las cualidades tímbricas de la voz.

Históricamente, los primeros estudios extensos para averiguar cómo se produce la voz, fueron realizados por Leonardo da Vinci (“Estudio de la Laringe”, 1490) y Andrés Vesalio (“De humani corporis fabrica”, 1543). En el siglo XVII, Girolano Fabricius d’Aquapendente, estudia la capacidad de la laringe para producir no sólo el habla, sino también “la maravillosa potencialidad y expresividad de la voz cantada”. Sus obras más importantes son “De Laryngis Vocis Instrumento” (1660); “De Locutione et ejus instrumentis” (1601) y “De Brotorum loquela” (1603). Claude Perrault publica en 1680 “De la musique des anciens”, donde explica la generación de las distintas tonalidades de la voz humana por las variaciones de longitud y de tensión de los pliegues vocales. En 1741, Antoine Ferrein, en su tratado “Sur l’organe immédiat de la voix et de ses différens tons”, llega a la conclusión de que en la laringe existen unas estructuras comparables a cuerdas de instrumentos musicales.

En el año 1800, Henri Dutochet, en su obra “Nouvelle Théorie de la Voix”, señala la importancia del músculo tiroaritenosoideo en la modificación del tono y compara la laringe con las boquillas vibrantes de los instrumentos de viento. Félix Savart, en 1825, publica sus “Memorias sobre la voz humana”, donde describe sus experimentos hechos en tubos de paredes membranosas y húmedas y los sonidos producidos por ellos.

El profesor Johannes Peter Müller (1801-1858), publicó en 1837 sus experimentos que permitieron establecer las teorías de la fonación vigentes hoy en día. Utilizó un aparato denominado “compresorium”, con el que regulaba la compresión medial³ de las cuerdas vocales, primero sobre lengüetas membranosas elásticas y, posteriormente, sobre laringes aisladas obtenidas de animales y cadáveres humanos.

Las investigaciones de Müller dieron lugar a la integración de la teoría mioelástica de la fonación, la cual completó y difundió Janwillem van den Berg en 1958. Las bases de estas teorías siguen siendo plenamente vigentes y han permitido el desarrollo de la teoría muco-ondulatoria enunciada por Jorge Perelló en 1962, para explicar el comportamiento vibratorio de las cuerdas vocales y completar el conocimiento en relación con la teoría propuesta por Minoru Hirano en 1975, quien explica que el movimiento de las cuerdas vocales tiene una doble estructura vibratoria. Hirano también habla acerca de la función de los músculos laríngeos: crico-tiroideo (tiende a alargar las cuerdas vocales), interaritenosoideo (no modifica el aspecto vibratorio), crico-aritenosoideo posterior (separa las cuerdas vocales, las alarga y las eleva) y de los demás músculos de la laringe que intervienen en la producción de sonidos en los diferentes registros de la voz [García-Tapia, 1996], [Seidner y Wendler, 2004].

En 1981, Dejonckere expone los resultados de sus investigaciones, que aportan datos que estaban insuficientemente explicados por las otras teorías. Sus conceptos permiten describir y pensar con un nuevo punto de vista muchos de los fenómenos de la patología vocal, como infiltraciones edematosas, neoplasias, neoformaciones, parálisis, distonias y diquinesias funcionales. Los elementos fundamentales de su teoría son el concepto de que el generador del sonido vocal es un oscilador de amortiguamiento débil, accionado mediante una fuerza periódica extrínseca; y el concepto de la importancia determinante de la impedancia mecánica para la regulación de la frecuencia vibratoria; por lo que se le denomina teoría oscilo-impedancial de la vibración de las cuerdas vocales.

³ La condición de cierre de las cuerdas vocales depende de la compresión medial, que es la relación entre la presión del aire subglótico y la fuerza de cierre necesaria para vencer su resistencia.

Por otro lado, el desarrollo de investigaciones acerca del procesamiento digital de la voz, inició a mediados del siglo XX, al generar voz sintetizada. El primer sintetizador vocal fue VODER (Voice Operating Demonstrator) presentado por Homer Dudley en Nueva York en 1939. VODER fue inspirado por VOCODER (Voice Coder) desarrollado en los Laboratorios Bell a mediados de los años treinta. El original VOCODER era un dispositivo para analizar el habla con una variación de parámetros acústicos, que luego se convierte en un sintetizador para reconstruir una aproximación de la señal original [Lemmetty, 2008].

Los estudios científicos sobre la producción de la voz, han contribuido al desarrollo de investigaciones acerca de la voz cantada en específico. Ello, unido a la inclusión de algoritmos computacionales y herramientas tecnológicas, ha permitido conocer algunos indicadores de ejecución vocal y expresiva de los cantantes, observables a partir de la señal acústica (forma de onda) y del espectro de la voz cantada⁴, tales como: el ataque y el final de la nota; el timbre, determinado por la producción de armónicos; y el vibrato, como resultado de modulaciones en la estabilidad del tono [Capponi, 2004].

El análisis del espectro de la voz cantada ha sido estudiado por muchos investigadores en las últimas cuatro décadas. Sin embargo, destaca el trabajo de Johan Sundberg, investigador sueco, quien desde la década de 1970, se ha dedicado a estudiar a cantantes y ha hecho experimentos sobre todos los aspectos conocidos hasta hoy en día, acerca del espectro de la voz cantada y su relación con la técnica vocal. Ha publicado diversos artículos sobre la resonancia, las formantes⁵, el tracto vocal, la función de los pliegues vocales, la respiración y su función en la producción vocal, el vibrato, el canto coral, la percepción del canto, la expresión en el canto y las características acústicas de voces patológicas.

A partir de las investigaciones de Sundberg, y las de sus colaboradores y contemporáneos, se han determinado algunos principios de interacción acústica que pueden utilizar los cantantes para transformar la calidad vocal a partir de los elementos fisiológicos, y que son detectados en el espectro de la voz.

Sundberg menciona que los cantantes (profesionales, entrenados en el canto clásico) trabajan para cambiar cinco modos de resonancia acústica de su tracto vocal (formantes), en particular: dos formantes que determinan la articulación de las vocales y son afectados por los movimientos de la mandíbula y la lengua, respectivamente; y los

⁴ El espectro es un gráfico de frecuencia - intensidad de una señal u onda determinada.

⁵ Una formante es un pico de intensidad en el espectro de un sonido; es decir, una concentración de energía que se produce en una frecuencia determinada.

otros tres que colectivamente integran la medida de la *formante del cantante*⁶, y el *ring* en una voz cantada [Sundberg, 1987]. El *ring* es un rango de frecuencias en el cual los armónicos de la frecuencia fundamental del cantante se alinean mejor con los correspondientes picos de resonancia del tracto vocal, creando el sonido que se irradia con la máxima riqueza y potencia en los labios. Sundberg lo define como una especie de “silbido suave y brillante” que se puede percibir al escuchar atentamente la voz potente de los cantantes mejor entrenados.

También, Sundberg y Skoog en 1997, Austin en 2007 y el mismo Sundberg en 2008, han publicado sus investigaciones sobre cómo las cantantes experimentadas aumentan la apertura de la mandíbula para ajustar la primera formante con la frecuencia fundamental. Este fenómeno, llamado *sintonía de formante* ("formant tuning" en la literatura anglosajona [Sundberg 1974], [Miller et al, 1997]), permitiría, en ciertas condiciones, favorecer la vibración de las cuerdas vocales y así facilitar la producción de las notas agudas [Titze, 1988, 2008], teniendo, sin embargo, repercusiones adversas en la inteligibilidad de las vocales, que las cantantes deben aprender a superar.

Existen muchas coincidencias de diferentes investigadores en todo el mundo, con los resultados que ha obtenido Sundberg. Por ejemplo, experimentando acerca del vibrato de la voz cantada, diversas investigaciones acústicas y fisiológicas se han enfocado en los mecanismos de modulación de frecuencia. De la misma manera, las investigaciones han establecido que otro componente del vibrato es la modulación de amplitud, que ocurre pasivamente cuando la frecuencia fundamental del cantante y los armónicos coinciden con las cuestas ascendentes y descendentes en el espectro de la función de transferencia del tracto vocal para determinada configuración articulatoria [Dromey et al, 2009]. Según Horii y Hata (1988), este proceso ocurre encima de la laringe, en el tracto vocal, y lo han denominado “interacción resonancia-armónicos” o RHI por sus siglas en inglés.

Investigadores como Seashore (1932), Mason (1965), Shipp, Leanderson y Sundberg (1980) y Lebon (1999), han realizado numerosos experimentos y estudios sobre

⁶ La investigación extensa de Sundberg en la producción vocal de cantantes de ópera lo condujo al concepto de "formante del cantante", una extra-resonancia asociada con el canto formal que no se presenta en el sonido hablado y que se obtiene cuando la laringe es bajada de tal modo que se amplía la base de la faringe.

el vibrato. Todos ellos coinciden en que el vibrato de los cantantes sumamente entrenados tiene una periodicidad promedio de 5 a 7 modulaciones por segundo⁷.

En 1994, los investigadores Robison, Bounous y Bailey llevaron a cabo un estudio con cantantes profesionales hombres y mujeres. Se realizó previamente una selección de las mejores voces por expertos en el área. Las voces elegidas fueron analizadas acústicamente para determinar cuáles propiedades eran las más esenciales a la belleza vocal. Las voces de más alto rango tenían muchos rasgos similares, principalmente el vibrato, con una medida promedio de 5.4 Hz. Su vibrato estuvo presente al menos durante el 75 % del tiempo cantante. Entre los factores respiratorios que contribuyen a la belleza vocal, predominó el uso de la respiración ventricular (utilizando principalmente el diafragma y músculos abdominales) sobre la respiración costal (usando los músculos de la caja torácica); así como la fuerza abdominal presente al inicio y en la duración de un tono. Además, las voces más hermosas tenían un timbre característico, que los autores definieron como "las fuerzas comparativas de la formante del cantante, la formante de vocal y el tono fundamental " o en otras palabras, las amplitudes proporcionales de estas formantes en los labios [Robison et al., 1994].

1.3 Modelo físico y análisis acústico de la señal de voz

La señal de voz es una onda acústica de presión sonora que se origina a partir de movimientos voluntarios del sistema físico de producción de la voz, que está conformado por los órganos anatómicos cuya fisiología corresponde a los procesos de respiración, fonación y resonancia [Bunch, 1997].

La producción de la voz ha sido integrada en un modelo físico, ampliamente aceptado por la comunidad científica: la denominada *teoría de la fuente y el filtro* (figura 1.3). Aunque su origen es anterior, el desarrollo matemático se debe al sueco Gunnar Fant (1960) [Rabiner, 1978]. La materia prima acústica se crea en la *fuentes*, o punto donde se genera el sonido básico que luego se habrá de filtrar o modelar. Esta fuente se sitúa en las cuerdas vocales si se trata de sonidos sonoros como las vocales y consonantes sonoras. En el caso de las consonantes sordas, la fuente está en el punto bucal donde se produce el ruido consonántico.

⁷ Sólo como dato histórico: dos grandes tenores, Caruso y Pavarotti, tenían una medida de vibrato de 6.5 Hz y 5.5 Hz, respectivamente (Dejonckere, 1995; Titze, Story, Smith y Long, 2002).

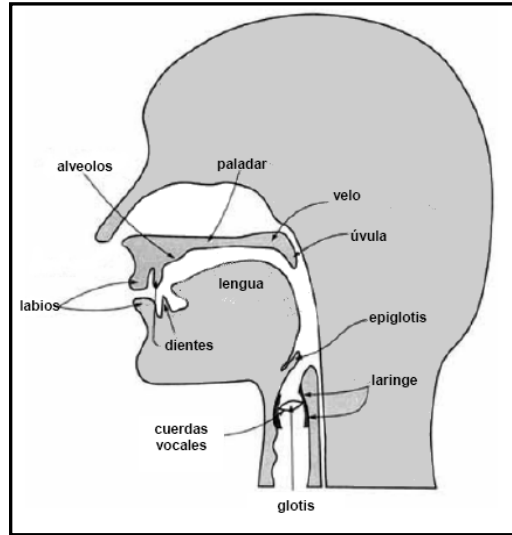


Figura 1.3 Fuente y filtro de la señal de voz.

La presión del aire que proviene de los pulmones, aumenta cuando está debajo de la glotis y ésta se encuentra cerrada por acción de la musculatura laríngea. Cuando es suficientemente grande esta presión por debajo de la glotis, llamada subglótica, provoca la apertura de la glotis y el aire pasa hacia arriba, haciendo aumentar la presión por sobre la glotis, denominada presión supraglótica.

A medida que la glotis comienza a cerrarse, la presión subglótica aumenta de nuevo. El flujo de aire procedente de los pulmones aumenta su velocidad intentando mantener el mismo paso de volumen, pero la inercia del aire al pasar por la laringe provoca un descenso muy rápido de la presión transglótica, debido al fenómeno conocido como efecto Bernoulli. Al mismo tiempo, la presión transglótica ayuda al súbito cierre de los repliegues vocales, que tiene como consecuencia la brusca interrupción del flujo de aire.

Así, la presión supraglótica, al interactuar con la fuente, ayuda a mantener la glotis abierta en la fase de apertura y a cerrarla de forma rápida y fuerte en la fase de cierre.

La continua alternancia de las fases abierta y cerrada de la glotis, libera pulsos de aire sucesivos, los cuales dan origen a la onda de sonido que pasará por el tracto vocal hacia los labios (figura 1.3.1) [Lindblom, 2007].

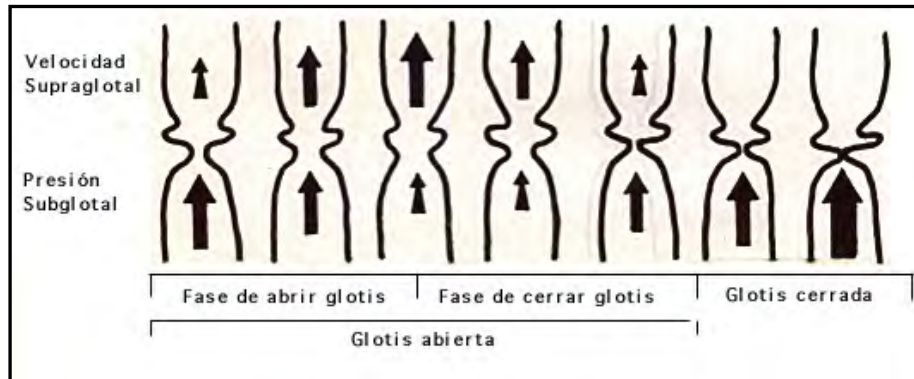


Figura 1.3.1 Ciclo glotal

Durante los periodos estables o estacionarios de los sonidos de tipo vocal, la onda generada repite el mismo ciclo a intervalos regulares de tiempo, por lo que es una onda periódica y cumple el Teorema de Fourier: está formada por una serie de armónicos (figura 1.3.2) [Fant, 1997]. El primer armónico es la frecuencia fundamental F_0 , y corresponde al tono de la voz; el segundo tiene una frecuencia doble de la fundamental, el tercero, triple, y así sucesivamente.

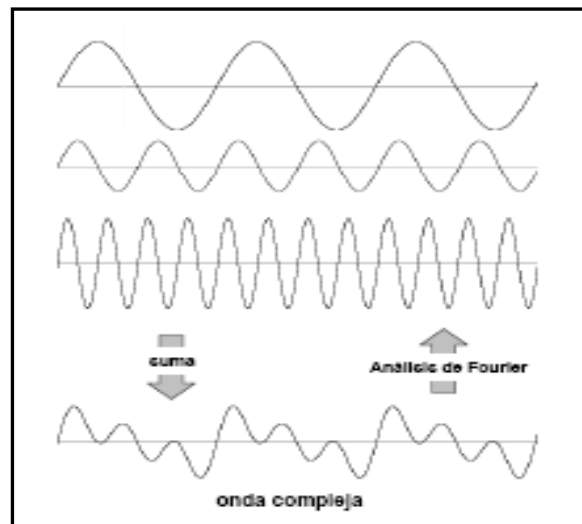


Figura 1.3.2. La onda compleja o periódica está integrada por armónicos de distintas frecuencias que se pueden determinar a través del análisis de Fourier.

El sonido generado en la fuente se propaga a través del tracto vocal y, de acuerdo con los principios de la *física de tubos* [Stanley, 1958], sufre un proceso complejo de *filtrado*, donde unas frecuencias se refuerzan y otras se atenúan.

Desde el punto de vista acústico, el tracto vocal de un varón medio, medido desde las cuerdas vocales hasta los labios, equivale a un tubo recto cilíndrico de 17.5 cm de largo (figura 1.3.3) [Chowning, 1981]. El tracto de la mujer es, en promedio, alrededor de 15% más corto que el del hombre. El hecho de que el tubo sea recto o curvado, como en el caso real, es irrelevante acústicamente. Lo importante es que el tubo experimente una constricción o estrechamiento en alguna parte de su recorrido. Dependiendo del lugar donde aparezca esta constricción, cambia completamente el conjunto de resonancias del tubo, dando lugar a resultados acústicos muy distintos (figura 1.3.4) [Vennard, 1967]. Al hablar o cantar se crean dichas constricciones debido a los movimientos de los órganos articulatorios, que son los labios, la mandíbula, la lengua y el velo del paladar (figura 1.3.5).

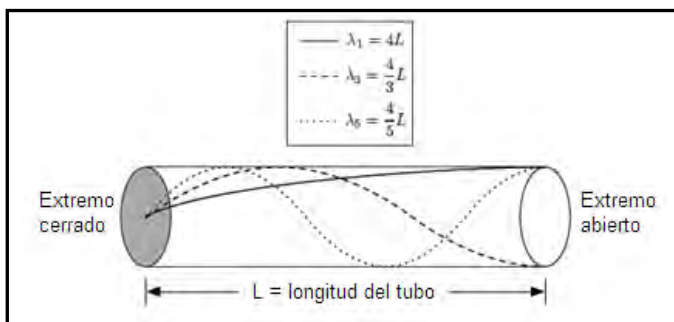


Figura 1.3.3 Tubo acústico que representa el tracto vocal, sin pérdidas.

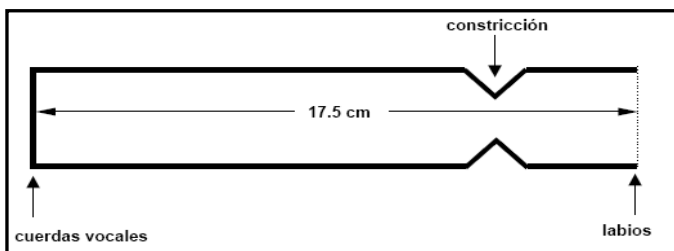


Figura 1.3.4 Tubo acústico que representa el tracto vocal, con una constricción que modificará sus resonancias.

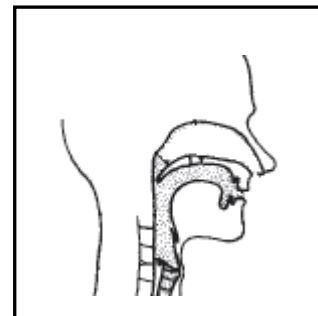


Figura 1.3.5 Representación esquemática del tracto vocal.

Los órganos articulatorios permiten concentrar la energía en determinadas frecuencias, actuando como resonadores. Estas frecuencias que se enfatizan, reciben el nombre de formantes (figura 1.3.6). Benade define a las formantes como los picos que se presentan en la envolvente espectral [Benade,1990]. Las frecuencias formantes más significativas del tracto vocal, en cuanto a la acústica de la voz hablada y cantada, son las primeras cinco (F1, F2, F3, F4 y F5). Dependiendo de la relación que se establezca entre las diversas formantes, es que se percibe un determinado timbre de voz o calidad vocal.

La frecuencia F0, que corresponde a la frecuencia fundamental de la voz, no se asocia con una formante, ya que no se origina como una resonancia del tracto vocal [Stanley, 1958].

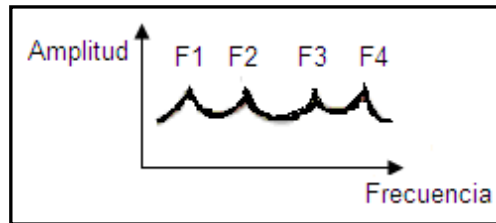


Figura 1.3.6 Curvas espectrales de frecuencias formantes.

El comportamiento de una formante, especificado en los dominios de tiempo y frecuencia, posee una amplitud y un ancho de banda. A menor ancho de banda, más selectiva es la formante para un determinado armónico, o grupo de armónicos cercanos, y será percibida más claramente. La curva que representa estas magnitudes está dada por una función de transferencia; es decir, representa la transmisión del sonido a través del tracto vocal (figura 1.3.7).

$$T(f) = p_b(f)/p_g(f)$$

Figura 1.3.7 Función de transferencia del tracto vocal.

$p_g(f)$ y $p_b(f)$ son las amplitudes espectrales de presión sonora en la glotis y en la boca, respectivamente.

Las dos primeras formantes (F1 y F2) permiten la identificación de las vocales. Las formantes agudas (F3, F4, F5) determinan el color de la voz.

En las vocales, la primera formante (F1) controla la amplitud del sonido y depende de la forma de la cavidad faríngea: cuanto más estrecha, mayor frecuencia y viceversa. La segunda formante (F2) controla la inteligibilidad del sonido y se supedita a la posición de la lengua: si esta se eleva en la parte anterior, la frecuencia subirá en relación directa con la altura y la anterioridad alcanzadas por la lengua; si es en la parte posterior, la frecuencia descenderá en relación inversa con la altura de la lengua. La tercera formante (F3) está relacionada con la acción de los labios. Su valor en frecuencia es más alto si éstos están estirados y más bajo si están redondeados. Las formantes F4 y F5 varían con

la anchura y longitud del tracto vocal; cuanto más corto y estrecho el tracto, más agudas estas formantes.

En general, alargar el tracto vocal por medio de la protrusión de los labios y bajar la laringe, hace bajar todas las formantes, oscureciendo el color de la voz. De modo inverso, acortar el tracto vocal al retraer las comisuras labiales y subir la laringe, eleva todas las formantes produciendo un timbre más claro [Bunch, 1997].

1.4 Parámetros acústicos característicos de la voz cantada

1.4.1. Frecuencia fundamental

La frecuencia fundamental F_0 corresponde a la frecuencia a la cual vibran las cuerdas vocales. Las propiedades de longitud, masa y elasticidad de las cuerdas vocales, son controladas por la musculatura laríngea. El acortamiento de la cuerda vocal provoca la disminución de la tensión y aumenta la masa de la superficie vibrátil, por lo que la frecuencia fundamental será disminuida. La elongación de la cuerda vocal produce un aumento en la frecuencia fundamental, al tensar la cuerda vocal y disminuir la masa y la superficie de contacto.

La presión subglótica afecta en cierto modo a la frecuencia fundamental. Al aumentar la presión subglótica, la frecuencia se incrementa por unos cuantos hertz (entre 2 y 7 Hz/cm H_2O) [Lindblom, 2007]. Por lo tanto, para que un cantante pueda mantener el tono durante un *crescendo*, necesita reducir la activación de los músculos laríngeos para regular la frecuencia fundamental de acuerdo con el incremento de la presión subglótica. Esto explica el hecho de que, quienes no han sido entrenados en el canto, no pueden cantar notas altas suavemente, ya que sus músculos laríngeos son incapaces de producir la tensión necesaria de las cuerdas vocales y, en lugar de ello, utilizan la presión de aire para elevar el tono.

Un cambio de la *aducción* glotal (la fuerza por la cual las cuerdas vocales se presionan una con otra) afecta el timbre de voz con respecto al modo de fonación. Un incremento de aducción lleva a una amplitud reducida de los pulsos de aire glotales y requiere una presión subglótica alta. Entonces, la amplitud relativa de la fuente de voz se reduce y la calidad de la voz se presenta como presionada, tensa o estrangulada (como al

levantar un objeto pesado). Por otro lado, con unas cuerdas vocales holgadamente aducidas, las cuerdas dejan de cerrar la glotis por completo, por lo que la forma de onda de la fuente de voz se vuelve casi sinusoidal y la voz resultante es como un susurro. Entre la voz susurrante y la fonación neural, hay otro tipo de voz, llamada flujo de fonación. Esta implica un flujo de aire generoso y la fuente de voz es fuerte. Típicamente se puede escuchar en cantantes de ópera de alto nivel.

La fuente de voz funciona de forma prácticamente independiente del tracto vocal, es decir, no es afectada por cambios en la articulación. Sin embargo, cantando en tonos muy altos, la frecuencia fundamental puede acercarse a la primera formante F1. En tales casos existe una fuerte interacción entre la fuente de voz y el tracto vocal. El flujo de aire transglotal es reducido o aún revertido durante un periodo corto de fase abierta, lo cual podría reducir el consumo de aire.

Los cantantes tienden a evitar la situación de que la frecuencia fundamental F0 sea más alta que la primera formante F1. Si la frecuencia fundamental fuese más grande que F1, la amplitud de la fundamental se reduciría al sobrepasar la frecuencia de máxima resonancia y el nivel de sonido radiado se reduciría también. En tonos muy altos las sopranos tienden a ampliar su quijada, casi sin considerar qué vocal están cantando, lo cual sugiere un cambio de articulación a un modo neutral de tracto abierto.

La adaptación de la forma del tracto vocal a la frecuencia fundamental afecta a todas las frecuencias formantes. Probablemente no sólo las sopranos aplican esta estrategia. Dependiendo de la frecuencia fundamental y del valor normal de F1 de la vocal cantada, la misma estrategia es aplicada también por barítonos, tenores y contraltos. Para vocales diferentes de la “a” probablemente se lleven a cabo cambios pequeños de la formante F1, mediante la disminución del grado de constricción de la lengua en el tracto vocal; pero para cambios mayores, los cantantes utilizan una ampliación de la apertura de la quijada. Esto sacrifica la inteligibilidad de la vocal cantada, en beneficio de la audibilidad [Sundberg, 1987].

La fuente de voz humana puede operar en diferentes modos de oscilación, que son los llamados *registros* vocales. De acuerdo a Titze, estos registros dependen de una interacción entre las resonancias subglotales y las vibraciones de las cuerdas vocales.

Rigurosamente, un cantante bajo es capaz de cantar aproximadamente a 80 Hz (tono E₂) y hasta 330 Hz (E₄), un tenor entre 123 Hz (C₃) y 520 Hz (C₅), una contralto entre 175 Hz (F₃) y 700 Hz (F₅), y una soprano entre 260 Hz (C₄) y 1300 Hz (E₆) [Lindblom, 2007].

1.4.2. Formante del cantante

Un cantante bien entrenado puede lograr conformar su tracto vocal para hacer que las formantes F3, F4 y F5 formen un grupo aglutinado estrechamente en frecuencia. A esta agrupación de formantes se le llama “formante del cantante”. Como las separaciones de frecuencia entre ellas disminuye, sus niveles individuales se incrementan y de ahí se obtiene un alto pico espectral entre 2500 y 3000 Hz. Esto es útil al cantar con el acompañamiento de una orquesta.

Por un lado, la formante del cantante se debe a un factor perceptual, ya que el oído humano goza de especial percepción a estas frecuencias y, por otro lado, destacan por encima del enorme sonido de una orquesta, pues la mayoría de los instrumentos no generan estas frecuencias con una amplitud suficiente. Una orquesta tiene un gran volumen de sonido entre los 500 Hz y los 1000 Hz. En esta región, los armónicos del cantante no pueden competir con la orquesta y resultan inaudibles; pero, gracias a una diferente distribución de energía en el espectro, el cantante acumula una intensa resonancia alrededor de los 3000 Hz. La existencia de una distancia constante entre armónicos, permite que el oído humano *reconstruya* las frecuencias enmascaradas que son inaudibles (F0 y sus primeros armónicos) y *perciba* la frecuencia fundamental, debido al fenómeno psicoacústico de percepción de la fundamental ausente o percepción de altura virtual [Roederer, 1995].

La situación acústica que produce el agrupamiento estrecho de F3, F4 y F5 es obtenido por enarmonías de frecuencia de la apertura del tubo laríngeo, también llamado tubo epilaríngeo, con la faringe. Esto puede lograrse debido al estrechamiento de esta apertura. Entonces el tubo laríngeo actúa como un resonador con una resonancia que no resulta muy afectada por el resto del tracto vocal, sino más bien por la forma del tubo laríngeo. Además del tamaño de la apertura y el tamaño del ventrículo laríngeo, también influyen el largo del ventrículo y la posición baja del tubo de resonancia laríngeo. Por lo visto, los cantantes templan el tubo de resonancia laríngeo a una frecuencia cerca de F3. El proceso articulatorio usado para llevar la cavidad a estas condiciones, parece consistir principalmente en un descenso de la laringe, que tiende a ensanchar a la propia laringe y al ventrículo laríngeo. Muchos maestros de canto recomiendan a sus estudiantes cantar con una confortable posición baja de la laringe.

El acercamiento de las formantes agudas en la formante del cantante se puede conseguir fundamentalmente de dos formas: haciendo descender la laringe, o bien, adelantando y elevando la lengua estrechando el vestíbulo laríngeo en sentido

anteroposterior, hasta formar un pequeño resonador que tenga una longitud aproximada de 1/6 de la longitud total del tracto vocal y su área tenga una sección de 1/6 del área de la faringe (figura 1.4.2) [Sundberg, 1987].

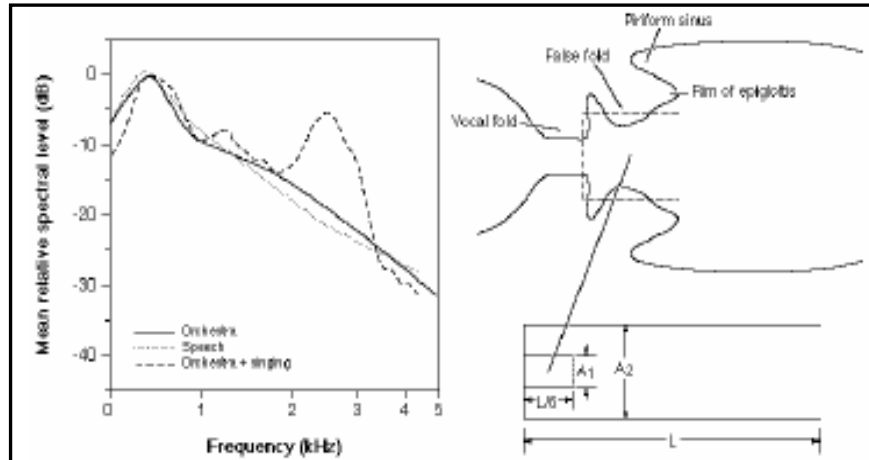


Figura 1.4.2 Acústica de la formante del cantante. Tomado de [Sundberg, 1987]

La frecuencia central de la formante del cantante varía ligeramente entre las diferentes clasificaciones de voz. El centro de frecuencia de la formante del cantante de bajos, barítonos y tenores está alrededor de 2.4, 2.6 y 2.8 kHz, respectivamente [Sundberg, 2001]. Estas pequeñas diferencias son bastante relevantes para los timbres típicos de voz de estas tesituras. Su origen es debido probablemente a diferencias en la longitud del tracto vocal; los bajos tienden a tener un tracto vocal más largo que los barítonos, quienes a su vez tienen un tracto vocal más largo que los tenores.

Es difícil que la formante del cantante aparezca en registros agudos de voz, como las sopranos. Esto se explica por el hecho de que cantan en una alta F_0 y tienen parciales extensamente espaciados en el espectro, por lo que resulta poco factible que alguno de estos parciales se ubique dentro del grupo de formantes F_3 , F_4 , F_5 , que es bastante estrecho en frecuencia cuando se produce la formante del cantante [Sundberg, 1987], [Benade, 1990].

Por lo tanto, se ha considerado que la formante del cantante es una característica común de cantantes de registros graves (bajos, barítonos, tenores, contraltos e incluso mezzosopranos) y entrenados profesionalmente. Sin embargo, Bloothoof y Plomp, en un estudio realizado en 1986 sugieren que, para determinar si existe formante del cantante en una vocal determinada, se puede considerar si su pico más alto se encuentra más allá

de -20 dB en comparación con el pico de la primera formante, y así, podría observarse este fenómeno en una voz de soprano. Sin embargo, los resultados de esta investigación no han sido muy ampliamente aceptados, pues otras investigaciones muestran que las sopranos utilizan el recurso de la sintonía de formantes con armónicos para producir los mismos efectos que la formante del cantante, como se explica más abajo [Bloothoof y Plomp, 1986].

Sundberg menciona que la formante del cantante no tiene fundamento en el canto no clásico, por ejemplo en el pop o en la música de teatro; donde la audibilidad es responsabilidad del ingeniero de audio, más que del cantante. De la misma manera, los cantantes de coro generalmente no articulan formante del cantante [Sundberg, 1987].

1.4.3. Sintonía formante - armónico

En el canto femenino de registro agudo, como es el caso de las voces de sopranos, se puede obtener un equivalente de formante del cantante, al lograr la llamada “voz resonante” que se obtiene abriendo mucho la mandíbula de forma que se sintoniza F1 con F0. También para notas en el registro medio, las formantes F1 o F2, o ambas, pueden ajustarse para coincidir ya no con la frecuencia fundamental de la voz F0, sino con alguno de los armónicos superiores; es decir, en todos estos casos ocurre la denominada *sintonización de formantes* (con armónicos). Las dos primeras formantes pueden ajustarse para sintonizar con los armónicos de la fuente, por lo que generalmente ya no guardan las mismas relaciones que en el habla. Este fenómeno puede no ser privativo de las sopranos, pues es probable que sea una tendencia natural de muchos cantantes profesionales de todas las tesituras, el sintonizar las frecuencias de los formantes con las frecuencias de los armónicos al producir la voz con mayor intensidad [Sundberg, 2000], [Benade, 1990].

1.4.4. Vibrato

La modulación, es decir, la variación cíclica sistematizada de un parámetro, como puede ser el de frecuencia y/o el de amplitud de la voz del cantante, constituye el *vibrato*. El mecanismo primario es la modulación de la frecuencia fundamental y secundariamente se ve afectada la amplitud. La frecuencia fundamental y sus armónicos varían cíclicamente, mientras que las formantes no experimentan cambios. La variación sincrónica con F0, hace que los armónicos se acerquen o alejen de las formantes, viéndose más o menos reforzados, de forma que la modulación en amplitud sobreviene

pasivamente cuando coinciden con las zonas de resonancia de la función de transferencia del tracto vocal [Seashore, 1932].

Cuando existe únicamente modulación de amplitud, en el canto se le denomina *trémolo*. Es considerado un defecto, pero ocurre en raros casos.

Si la modulación de la frecuencia fundamental es producida únicamente en la glotis, el vibrato tendrá una frecuencia menor a 5 Hz, que en inglés se denomina *bleat*. Es común que sea producido por cantantes sin entrenamiento o por cantantes jóvenes en los primeros años de su formación.

También es posible que el vibrato se genere desde el diafragma (por falta de control o tono muscular) y que tenga una frecuencia entre 6.5 y 8.0 Hz. Este tipo de vibrato frecuentemente afecta la claridad del canto. Es uno de los defectos en voces envejecidas. En inglés, se denomina *wobble*.

El sistema vocal de un cantante bien entrenado, posee una tendencia natural de producir un vibrato cuya frecuencia se encuentra entre 5 y 7 Hz [Sundberg, 2000], [Titze et al., 2002].

1.5 Programas existentes para el análisis de audio y voz

La industria informática del canto ha producido un gran número de aplicaciones y programas que utilizan los principios de la investigación técnica y científica sobre la voz, pero que están hechos con una finalidad comercial y de entretenimiento, tales como diversos tipos de sistemas de “karaoke” y transformadores de distintas características vocales, que cambian el timbre o la entonación. Este tipo de programas no serán tomados en cuenta en este trabajo.

En cambio, los programas de cómputo cuya utilidad se relaciona con el trabajo aquí mostrado, consideran dos campos de investigación: el análisis digital de señales acústicas y el estudio de la voz cantada.

En el **apéndice A** de esta tesis, se muestra una lista de los programas que existen en la actualidad, clasificados en tres grupos de acuerdo a su área de aplicación y al tipo de usuarios que pueden manipularlos. Con la finalidad de realizar una comparación de la funcionalidad de dichos programas con la del programa aquí desarrollado, se presenta a continuación un breve análisis de su funcionamiento, bajo el enfoque de su posible uso por un cantante.

1. Programas para análisis acústico de señales de audio.

Estos programas no se enfocan específicamente a la voz, pero en ellos se presenta el espectrograma del sonido que se alimente, ya sea por medio de un micrófono o de un archivo de audio digital determinado.

El programa *Acoustic Analyzing System* es considerado de alta especialización científica, pues está dirigido a expertos en Acústica y en Procesamiento Digital de Señales. En realidad es un módulo que junto con otros dos: un *analizador en tiempo real*, y un *analizador ambiental de ruido*, constituye un “laboratorio” acústico. Las funciones de este laboratorio se pueden aplicar en mediciones electrónicas de audio, acústica de espacios y análisis de: ruido, voz humana e instrumentos musicales; así como en el campo de la “telemedicina” para realizar auscultaciones sonoras. El módulo analizador acústico contiene un generador de señales de audio y un analizador FFT que se puede utilizar para realizar mediciones y correcciones de frecuencias, por ejemplo para afinar algún instrumento musical, entre otras aplicaciones. Con el espectrograma que este programa puede generar, se determina la frecuencia fundamental, las frecuencias de armónicos y de picos formantes. Siendo un programa muy completo, su interfaz gráfica contiene muchos elementos numéricos y matemáticos relativos a términos técnicos de la Acústica, cuyo uso y comprensión resulta inadecuado para un músico artista.

Los programas *Audio Sculpt* y *Wave Surfer* son representativos del tipo de programas de cómputo que se pueden utilizar en un estudio de grabación, pues con ellos se puede grabar y editar una señal de sonido; es decir, cambiar digitalmente sus características para realizar correcciones en tiempo y frecuencia. También muestran espectrogramas para determinar visualmente el análisis FFT. Es necesario tener conocimientos en procesamiento digital de señales para poder utilizar cómodamente todas las funciones de estos programas.

2. Programas para análisis médico de la voz.

Los programas *Computerized Speech Lab* y *Dr. Speech*, son aplicados principalmente en el campo de la medicina, para la prevención y la detección de patologías vocales. Generalmente tratan de la voz hablada, pero pueden aplicarse para estudios a cantantes. Presentan electroglotogramas y espectrogramas, de los cuales se debe hacer una interpretación basada en conocimientos especializados sobre fonoaudiología o foniatría.

Los programas *Voxmetria* y *Voce Vista* tienen la intención de servir también al análisis acústico de la voz cantada, pero también desde el punto de vista del tratamiento de las enfermedades vocales. Sus espectrogramas deben ser interpretados visualmente por un especialista.

3. Programas para entrenamiento de la voz cantada.

Singing Tutor, sugiere por su nombre, que fue creado con la finalidad de proporcionar al cantante una manera de entrenar su voz. Su función principal es la detección de altura de la nota musical cantada, enfocada a la práctica de la entonación de melodías.

El programa *Sing and see*, en su versión “simple” no trata otras características vocales más allá de la entonación. Ya en su versión “profesional” incluye un espectrograma de la voz cantada, el cual debe ser interpretado por el profesor de canto.

Por último, el programa *Overtone Analyzer* es un programa muy reciente, que también presenta un espectrograma. En su interfaz gráfica incluye un teclado y los dibujos de una clave de sol y una clave de fa, pues fue creado con la finalidad de que sea utilizado por músicos y profesores de canto.

Después de revisar y analizar las características generales de los programas presentados, es posible notar que en ninguno se realiza una determinación completa de las características acústicas (incluyendo entonación, resonancia vocal y vibrato) a partir del análisis digital de la voz cantada, para su aplicación en la profesionalización del canto.

Existen múltiples programas de cómputo que no son descritos aquí y que son capaces de presentar el espectro de un determinado sonido. Sin embargo, tienen la característica común de que sus resultados no evalúan directamente los parámetros acústicos de la voz cantada, sino que para estimarlos de manera aproximada, dependen de la interpretación visual del usuario.

Por lo tanto, se puede afirmar que no existe, hasta el momento en que se escribe esta tesis, un programa de cómputo que reúna las características planteadas para *Evocanto*, el programa de cómputo desarrollado en este trabajo; pues *Evocanto*, además de mostrar el espectro de la voz, presenta los resultados del análisis en una forma clara y precisa, sin la necesidad de realizar interpretaciones técnicas especializadas. No se tiene referencia de otro programa que proporcione resultados de esa manera, de las cualidades tímbricas específicas de la voz cantada, para ser utilizados por un estudiante de canto.

CAPÍTULO 2. ELABORACIÓN DEL PROGRAMA INTERACTIVO “EVOCANTO”

Para elaborar el programa interactivo de cómputo *Evocanto*, fue necesario establecer las necesidades de diseño y desarrollo del mismo, para elegir un entorno de programación adecuado, incluyendo el lenguaje de programación y las bibliotecas de apoyo especializadas. Eso es lo que brevemente se describe en las primeras partes de este capítulo. Posteriormente se detalla el funcionamiento del programa *Evocanto* y el diseño de la interfaz gráfica.

2.1 Análisis de requerimientos

El programa *Evocanto* está dirigido principalmente a estudiantes de canto profesional y a profesores de la misma área, que deseen incorporarlo en sus clases o en su metodología de enseñanza.

El objetivo más importante del programa *Evocanto* es que permita realizar un análisis acústico comprensible de la voz cantada, siendo posible evaluar ciertas cualidades vocales. Para lo cual debe ser fácilmente manipulable por el usuario y servir como herramienta complementaria en el estudio y comprensión de la técnica vocal.

Una vez activado el funcionamiento del programa, y con el uso de un micrófono para captar la señal de voz, el usuario cantará una nota, con una vocal, sosteniéndola durante algunos segundos. Al mismo tiempo, en el programa *Evocanto* serán llevados a cabo los siguientes procesos:

- Análisis acústico de la voz cantada, que consiste en la obtención del espectro de la señal de voz y, a partir de él, la detección de armónicos y la envolvente espectral; la determinación de la frecuencia fundamental F_0 y de las frecuencias formantes F_1, F_2, \dots ; así como la amplitud o nivel de energía.
- Verificación de la existencia de formante del cantante y del grado de sintonización de formantes con armónicos.
- Evaluación de entonación y vibrato, determinando la información correspondiente a la nota musical cantada y a la cantidad o frecuencia de vibrato.
- Determinación aproximada de la forma o configuración del tracto vocal, de acuerdo a la relación existente entre las primeras dos formantes.

Es deseable que el manejo de *Evocanto* sea cómodo e interesante. Esto se logra al existir interactividad entre el usuario y el programa, por medio de la interfaz gráfica [Gándara, 1994]. Por ello, *Evocanto*, además de poder ser usado con facilidad operacional, proporcionará una gran interactividad visual, para que, así, el usuario pueda verificar lo que está ocurriendo con su voz en el momento de cantar. Esto proporcionará al usuario la facultad de experimentar cambios en la postura corporal, en la colocación de la voz, los movimientos de labios y mandíbula; así como ajustes en la entonación y el vibrato de la nota cantada.

2.2 Herramientas de programación

De acuerdo con los requerimientos del programa, para desarrollarlo, surgió la necesidad de encontrar un lenguaje de programación flexible y eficiente, con alto nivel de abstracción, en el cual se disponga un manejo de señales de audio y análisis espectral, y que cuente con herramientas para desarrollo de la interfaz gráfica.

Con base en su eficiencia y flexibilidad, fueron seleccionadas las siguientes bibliotecas multiplataforma de código libre, desarrolladas en C++:

- SndObj, para la adquisición de audio y análisis espectral, y
- WxWidgets, para la interfaz gráfica

Sin embargo se optó por no utilizar directamente el lenguaje C++, sino el lenguaje Python, porque ofrece más ventajas y se puede utilizar de modo interactivo, lo que facilita hacer experimentos durante el desarrollo del programa. Además de que posee una sintaxis sencilla, clara y elegante [Van Rossum, 2003].

2.2.1 Python

Python es un lenguaje de programación creado por el holandés Guido Van Rossum en el Centro de Investigación Nacional de Matemáticas e Informática, de Países Bajos (CWI), en 1990. El nombre del lenguaje proviene de la afición de Van Rossum, hacia el grupo de humoristas británicos “Monty Python” [1].

Actualmente, Python pertenece a la *Python Software Foundation* (PSF), una organización sin ánimo de lucro, creada en el año 2001 en Estados Unidos, para la promoción y la protección de Python como lenguaje de código abierto. En febrero de 2002, Van Rossum fue premiado con el FSF Award for the Advancement of Free Software.

Las principales características del lenguaje Python, son las siguientes:

- Multiplataforma. Puede ejecutarse bajo cualquier sistema operativo de uso común.
- Multiparadigma. Además de la programación orientada a objetos, admite la programación imperativa, programación funcional y programación orientada a aspectos.
- Modular. Permite dividir el programa en módulos, también reutilizables desde otros programas Python.
- Interpretado o de script. No es necesario compilar el código a lenguaje máquina. El código fuente se traduce a un pseudo código máquina llamado *bytecode* la primera vez que se ejecuta, generando archivos *.pyc* o *.pyo* (*bytecode* optimizado), que se podrán ejecutar en sucesivas ocasiones.

Python ofrece varias bibliotecas de clases (o *toolkits*) para la implementación de una aplicación gráfica. Entre otras, las más conocidas son: Tkinter, PyGTK, PyQt, Pythonwin y wxPython.

Tkinter es una adaptación en Python del módulo de interfaz gráfica Tk ofrecido por el lenguaje de programación Tcl. Se distingue por ser multiplataforma y venir integrado con la distribución estándar de Python. Sin embargo, no presenta una apariencia nativa, es decir, que independientemente de la plataforma (Windows, UNIX o Mac), ofrece por defecto, la misma apariencia. Su uso es muy sencillo, pero ofrece un número limitado de controles gráficos.

Las bibliotecas GTK y Qt se constituyen en módulos para su uso en Python, y son llamados PyGTK [3] y PyQt [2]. También existe Pythonwin [4]. Están orientados a una plataforma específica (UNIX o Windows), aunque las interfaces que producen son muy rápidas, dado que interactúan directamente con funciones de interfaz gráfica nativas de la plataforma. Su uso es más complicado que en el caso de Tkinter.

El módulo wxPython [5] es multiplataforma, presenta una apariencia visual nativa y ofrece un alto rendimiento gráfico. Está basado en la *Interfaz de Programación de Aplicaciones* (API) wxWidgets, para C++. Permite realizar interfaces gráficas muy sofisticadas. Combina eficiencia con sencillez de uso.

La API wxWidgets fue creada para proveer una manera eficiente y flexible de maximizar la inversión realizada en el desarrollo de aplicaciones con interfaz gráfica. A pesar de que otras bibliotecas de clases ya existían para desarrollo multiplataforma,

ninguna cumplía todos los criterios deseables de bajo precio, código abierto, facilidad de programación y soporte de un amplio rango de compiladores.

Hasta febrero de 2004, wxWidgets era conocida como wxWindows. Sin embargo, dadas las repercusiones legales que el uso del término “Windows” podría tener, ya que es una marca registrada por Microsoft, los creadores de wxWindows y wxPython, Julian Smart y Robin Dunn, respectivamente, decidieron, renombrar el proyecto como wxWidgets.

El módulo wxPython añade una capa de programación simple de Python encima de wxWidgets, aislando la complejidad de la programación en C++ y ofreciendo aún la misma funcionalidad de esta completa biblioteca.

2.2.2 pySndObj

La biblioteca de objetos de sonido SndObj [6] surgió en 1997 como un proyecto de investigación en la Universidade Estadual, en el sur de Brasil, financiado por el Consejo de Investigación brasileño CNPq. El proyecto fue desarrollado en el Núcleo de Música Contemporánea, por Víctor Lazzarini, asistido por Fernando Accorsi.

SndObj es una biblioteca orientada a objetos. Es una colección de más de cien clases para la síntesis y el procesamiento del audio digital. El código fuente es multiplataforma y puede ser utilizado en C++, Java o Python, donde es insertado como un módulo. En el caso de Python, este módulo es llamado pySndObj [Lazzarini, 2000].

Las características más importantes de esta biblioteca son:

- Encapsulación: las clases de la biblioteca encapsulan todos los procesos complicados de producción, control, manipulación, almacenaje y funcionamiento de datos de audio.
- Modularidad: los objetos pueden ser asociados libremente, como si fuesen módulos o generadores, cada uno realizando funciones específicas.
- Portabilidad: Permite la especialización dependiente de máquina cuando es necesario, es decir, cuando el funcionamiento de los objetos de audio es específico de la plataforma.

2.3 Diseño y desarrollo del programa.

El programa *Evocanto* está conformado por tres procesos principales, por medio de los cuales se digitaliza la señal de audio, se analiza espectralmente a ésta y se presentan los resultados en forma gráfica (figura 2.3).

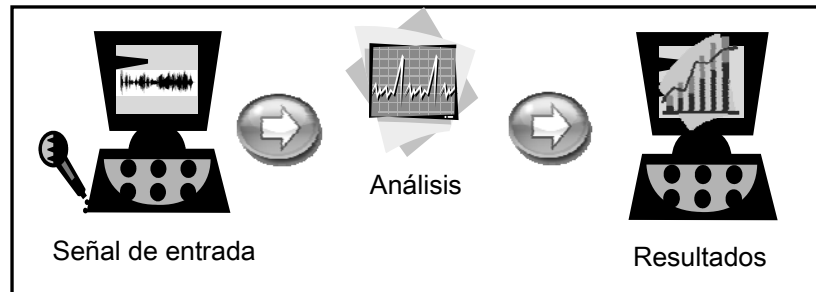


Figura 2.3 Procesos principales de *Evocanto*

2.3.1 Conformación de la señal de entrada.

La configuración típica de la entrada de audio consta de un micrófono y una interfaz de audio o tarjeta de sonido. Puede ser utilizada la propia tarjeta de sonido de la computadora y un micrófono convencional para computadora. Sin embargo, esta configuración tiene el inconveniente de captar ruido electrónico proveniente de la misma computadora, que interfiere en el buen funcionamiento del programa, por lo que es recomendable utilizar una interfaz de audio externa y, de esta manera, poder utilizar también un micrófono de mayor calidad, con mejor captación de la señal de voz.

Para digitalizar y obtener el espectro de la señal de entrada, *Evocanto* utiliza varias clases (objetos computacionales) de la biblioteca SndObj [Lazzarini, 2000]. Los objetos de la clase SndRTIO y SndIn, dan acceso a la interfaz de audio definida por omisión o predeterminada en el sistema operativo, con referencia específica a la señal de entrada de audio número 1. Un objeto de la clase HammingWindow se utiliza para definir una ventana de ponderación tipo Hanning, que se aplicará a cada bloque de muestras de la señal de entrada de audio. Finalmente, un objeto de la clase FFT implementa la *Transformada Rápida de Fourier*. Los objetos indicados previamente, se enlazan en una secuencia o hilo de procesamiento, utilizando un objeto de la clase SndThread, e intervienen en la clase FFT para transformar cada bloque de muestras de la señal de tiempo, en un espectro de frecuencia.

La señal de entrada es determinada por el tamaño de bloque que se va tomando de la interfaz de audio, luego es *ventaneada*, transformada por la FFT y escalada. Este proceso ocurre a intervalos regulares, determinados por el tamaño especificado para el tamaño del bloque de muestras de la entrada de audio. La salida proporciona un vector con datos de la FFT, conteniendo un par real-imaginario para cada punto de frecuencia sobre el lado positivo del espectro.

En *Evocanto* se utiliza una frecuencia de muestreo f_s de 44100 hertz, debido a que este es un valor estandarizado muy común en las aplicaciones de audio digital en las computadoras actuales.

El tamaño de la FFT se elige normalmente, por razones de mayor eficiencia computacional, como una potencia de 2. Debido a ello, y para proporcionar el mejor equilibrio entre la duración de cada bloque de audio y la resolución de frecuencia en el análisis espectral de la señal, se estableció el tamaño de bloque de audio en $N = 32768$.

La duración del bloque de audio y la resolución de frecuencia se determinan como se explica en seguida.

La resolución temporal dt , también llamado periodo de muestreo, corresponde a:

$$dt = 1 / f_s;$$

que en este caso equivale a $dt = 1 / 44100 \text{ Hz} = 0.000,023 \text{ s} = 23 \mu\text{s}$ (23 microsegundos, o 23 millonésimas de segundo), que es el tiempo que transcurre entre cada muestra.

La resolución espectral df , corresponde a:

$$df = f_s / N;$$

donde N es el tamaño de bloque, es decir, el número de muestras. Por lo tanto, el equivalente es: $df = 44100 \text{ Hz} / 32768 = 1.3 \text{ Hz}$, que es el intervalo de frecuencia entre los puntos del espectro.

La duración del bloque de audio, puede ser calculada en segundos mediante los siguientes cálculos:

$$T = N dt = N / f_s = 1 / df$$

De acuerdo a ello: $T = 32768 / 44100 \text{ Hz} = 0.74 \text{ s}$.

Una vez establecida la captura por bloques de la señal de audio, esta pasa a través de una ventana de ponderación de amplitud en el tiempo, que en *Evocanto* es de tipo Hanning. Esta ponderación se utiliza normalmente en el análisis digital de señales para empalmar y suavizar los cortes entre las ventanas sucesivas de análisis de la FFT, es decir, pondera a cero los extremos de las mismas (figura 2.3.1).

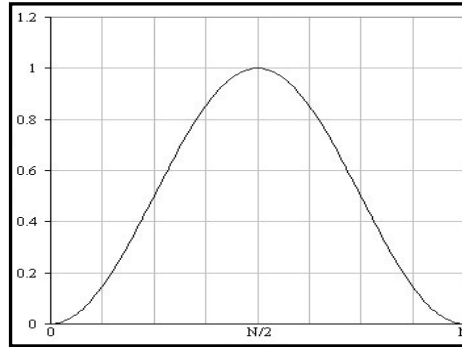


Figura 2.3.1 Ventana Hanning

La máxima frecuencia que se muestra en el espectro de *Evocanto*, se estableció en 5000 hertz, debido a que este límite es suficiente para incluir las formantes superiores más relevantes en el análisis de la voz.

Cada componente del espectro está conformado por un número complejo representado por un punto con una coordenada x (parte real) y una coordenada y (parte imaginaria). Su correspondiente magnitud cuadrática, normalizada al tamaño de bloque, es calculada mediante la fórmula:

$$S^2 = \frac{x^2 + y^2}{N}$$

donde S es la magnitud y N el tamaño de bloque de la señal.

Esta magnitud cuadrática en unidades lineales; es decir, no logarítmicas como en el caso de los decibeles, se utiliza posteriormente para calcular la frecuencia de vibrato. En seguida convierte a decibeles (NPS: nivel de presión sonora), mediante la fórmula:

$$\text{NPS} = 10 \log_{10}(S^2)$$

para poder conformar el espectro en unidades logarítmicas (decibeles). En *Evocanto* no es necesario que el nivel de presión sonora se calibre en unidades físicas.

2.3.2 Análisis espectral de la señal de voz cantada.

Una vez obtenido el espectro, se detectan los picos espectrales, los cuales corresponden a máximos en el rango de frecuencia con un umbral de magnitud. Estos se utilizan para realizar una estimación de la frecuencia fundamental F_0 , de manera aproximada, identificando la serie armónica que mejor explica los picos espectrales observados. A continuación se mejora esa estimación inicial calculando la frecuencia

promedio, ponderada por las magnitudes espectrales. Es decir, se obtiene el centroide de frecuencia [Potter, 1950], con la siguiente fórmula:

$$F0 = \frac{\sum w_i f_i}{\sum w_i}$$

donde f_i es la frecuencia del i -ésimo componente; w_i es un factor que indica la cantidad de decibels que el i -ésimo componente está debajo del máximo y equivale a la relación S_i/S_0 , entre amplitudes, expresada en decibels.

La frecuencia fundamental calculada se obtiene en hertz y entonces se calcula su equivalencia en términos de notas musicales, así como la diferencia en cents c , en relación con el La_5 de 440 Hz, basando los cálculos en la siguiente fórmula:

$$c = 1200 \log_2 \left(\frac{F0}{440 \text{ Hz}} \right)$$

Para determinar la nota, el índice de octava y la desviación en cents, se procede como en el siguiente ejemplo:

Si $F0 = 600$ Hz, entonces $c = 537$ cents por arriba del La_5 de 440 Hz. Este valor en cents se expresa como un múltiplo de 100, en este caso 500 cents (cinco semitonos), y un residuo de 37 cents. Por lo tanto, la nota resultante está cinco semitonos arriba del La_5 , con una desviación de 37 cents. Esto corresponde a la nota Re_6 con una desviación de +37 cents. En este caso, el índice de octava se incrementa de 5 a 6.

Un procedimiento general que equivale al ejemplo mostrado, se encuentra programado en el código de *Evocanto*.

La frecuencia de vibrato de $F0$, es determinada utilizando un método indirecto, se toma en cuenta la anchura espectral como un valor aproximado del vibrato, sin distinguir modulaciones de amplitud o modulaciones de frecuencia. La anchura espectral se calcula mediante la fórmula [Cordourier y Orduña, 2009]:

$$\Delta f = \sqrt{\frac{\sum_{n=N_1}^{N_2} [f_n - f_c]^2 \cdot S(f_n)}{\sum_{n=N_1}^{N_2} S(f_n)}}$$

Si el resultado obtenido es menor de 5 Hz, se considera que la cantidad de vibrato es menor de lo normal (para un cantante entrenado). Si es mayor de 7 Hz es exagerado; y si se encuentra en el intermedio, es considerado normal [Sundberg, 2000], [Titze et al., 2002].

Se ubican los armónicos, primero en forma aproximada, tomando múltiplos de la frecuencia fundamental F_0 . Posteriormente se mejora la aproximación, determinando los máximos picos espectrales más cercanos a estos múltiplos.

Para localizar los picos de las frecuencias formantes F_1, F_2, \dots, F_n , se utiliza un proceso de detección de máximos y mínimos en las componentes armónicas del espectro. Se comparan progresivamente las amplitudes de los armónicos para determinar si la envolvente espectral sube o baja. En los puntos de máxima amplitud se localizan las formantes. Se estableció un valor mínimo de 200 Hz, para considerar una formante.

El algoritmo utilizado se muestra en la siguiente figura (2.3.2).

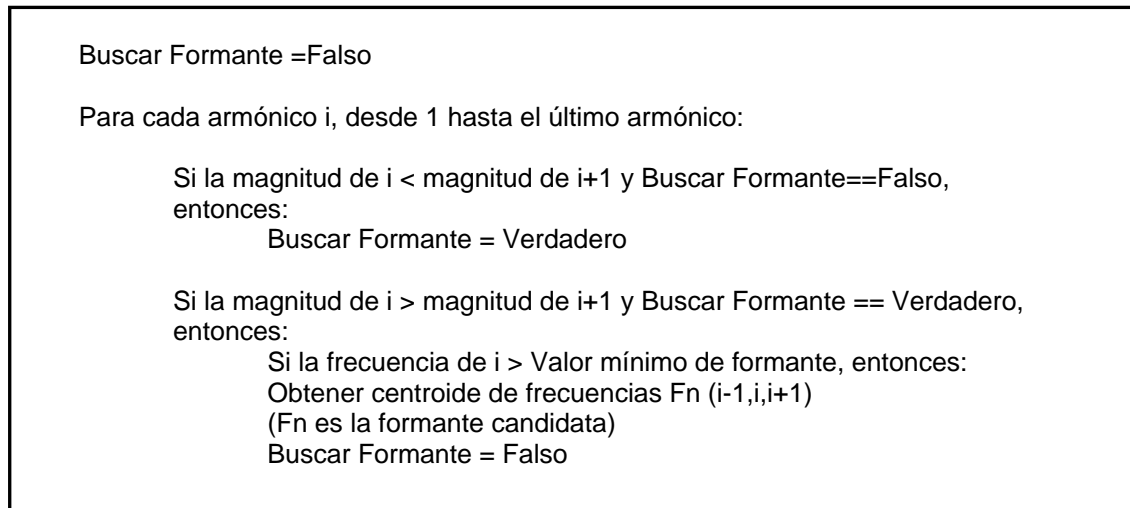


Figura 2.3.2 Algoritmo de *Evocanto* para localizar frecuencias formantes en el espectro.

El centroide de frecuencias de armónicos para F_1 (similar para F_2, F_3 , etc.), se obtiene con la siguiente fórmula:

$$F_1 = \frac{\sum_{i=1}^{i=1} f_i \cdot S(f_i)}{\sum_{i=1}^{i=1} S(f_i)}$$

f_i corresponde a la frecuencia y $S(f_i)$ a la magnitud, de cada armónico.

La sintonía de las formantes F_1 y F_2 con armónicos, se establece de la siguiente manera, para F_1 :

$$\Delta F_1 = (F_1 - F_i) / F_0 * 100\%$$

Con una fórmula similar para F_2 .

Es decir, equivale a la diferencia de frecuencias entre la formante correspondiente y el armónico alto F_i más cercano a F_1 . Se multiplica por 100% para calcular la sintonía

formante – armónico como el porcentaje de desviación o desajuste, entre -50% y +50%; los signos negativo (-) o positivo (+), indican, respectivamente, que la formante está debajo o encima del armónico.

El criterio utilizado para determinar si existe formante cantante, es constituido por dos valores importantes. El primero es la frecuencia fundamental máxima, establecida en *Evocanto* en 300 Hz. El segundo es la tercera formante F3, que debe ser mayor que 2400 Hz. Esto se debe a que sólo puede existir formante cantante en registros graves de voces graves, pues en notas de registros agudos las formantes están más espaciadas [Sundberg, 1987], [Benade, 1990].

Con base en los parámetros de las formantes F1 y F2, se establecen los criterios de articulación vocal. F1 indica el grado de apertura de la boca y F2 el grado de retracción del cuerpo de la lengua. En *Evocanto* se establecen las regiones vocálicas, de acuerdo a la siguiente gráfica de vocales para el idioma español (figura 2.3.3) [Quilis y Esgueva, 1983].

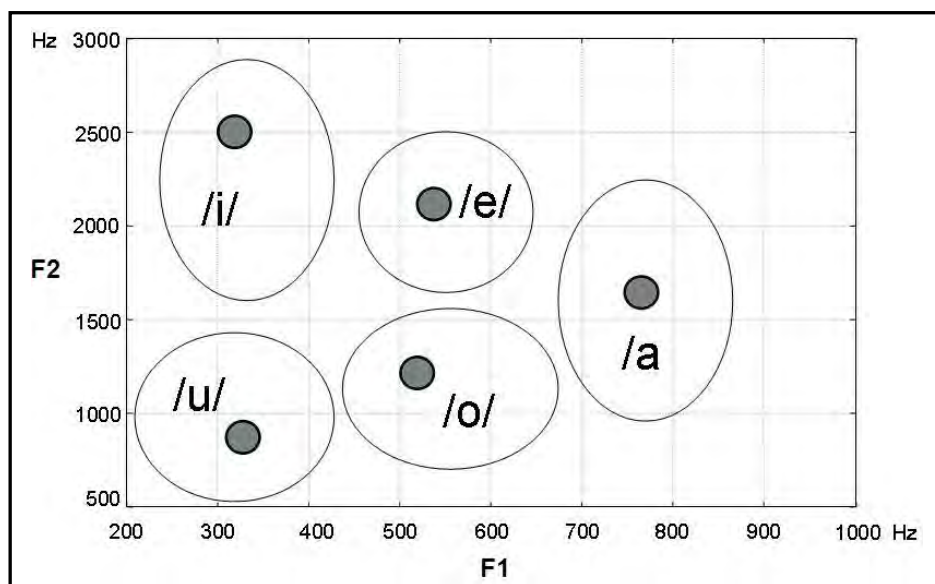


Figura 2.3.3 Regiones vocálicas del idioma español.

Sin embargo, en *Evocanto*, la relación entre las formantes F1 y F2 es utilizada como una aproximación para determinar la forma o configuración general del tracto vocal y no se indica la vocal correspondiente. Esto se debe a que intervienen diversos factores adicionales, incluso perceptuales, para distinguir y establecer con precisión cualquier vocal [Potter, 1950], [Ladefoged, 1989], [Neary, 1989].

2.3.3 Presentación de los resultados.

Cuando el usuario ha configurado la entrada de audio de manera que corresponda a la señal (normalmente de un micrófono) que se desea analizar en *Evocanto*, entonces puede dar click en el botón de encender, para realizar el análisis de la voz cantada. Se despliegan los resultados en la interfaz gráfica, con un retraso de 0.7 segundos del tiempo real, debido a que esa es la cantidad de tiempo que *Evocanto* ocupa en realizar el análisis.

Si el programa no detecta entrada de audio, envía un aviso en la barra de estado. El usuario tendrá que revisar la configuración de entrada de audio por medio de la interfaz de audio externa, o por micrófono convencional, según sea el caso.

Evocanto consta de una sola pantalla principal, donde se muestran los resultados de la siguiente manera (figura 2.3.4):

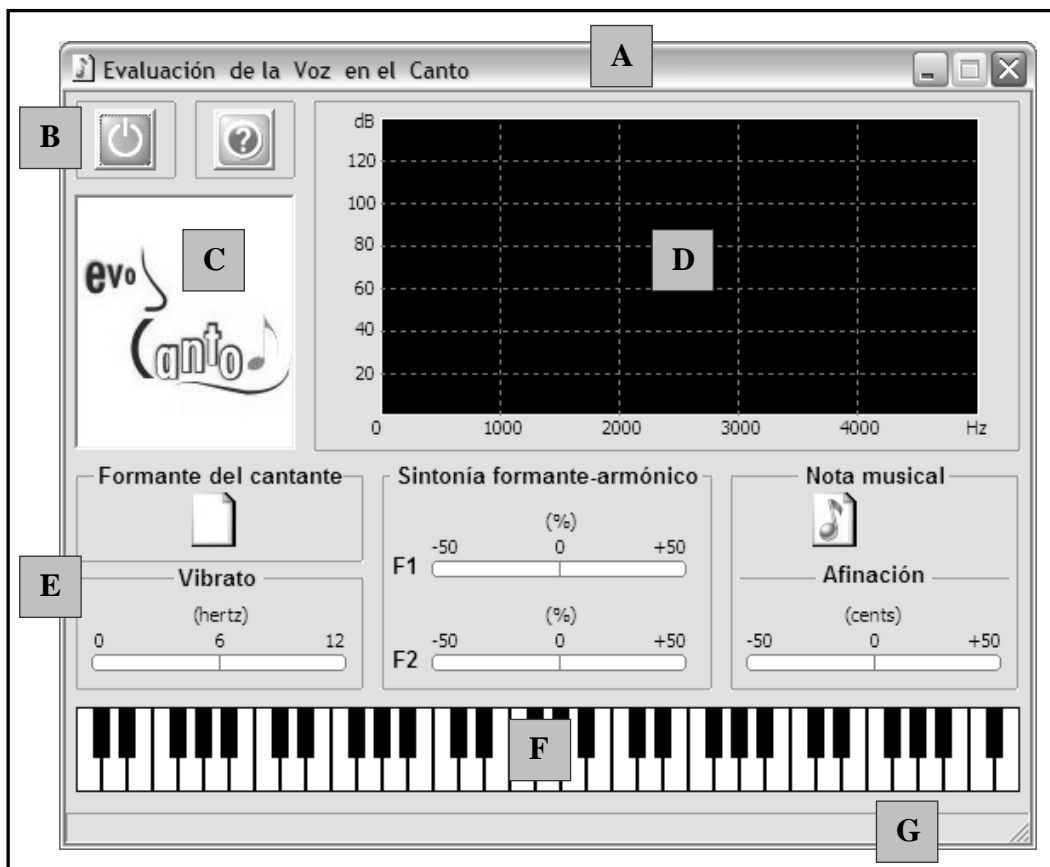


Figura 2.3.4 Elementos de la pantalla principal de *Evocanto*

A. Barra de icono, título y botones de minimizar y cerrar ventana.

B. Área de botones de control, con un botón de encendido/apagado para activar o desactivar el funcionamiento del análisis de la voz cantada y un botón para ayuda⁸.

C. Área de logotipo del programa y diagramas de Tracto Vocal.

En el área de logotipo, se muestra la imagen correspondiente a la forma o configuración general del tracto vocal, de acuerdo a la siguiente tabla 2.3, basada en el área de las regiones vocálicas de la anterior gráfica de la figura 2.3.3.

D. Área del espectro de la voz cantada.

El espectro de la señal de voz cantada, es integrado por la representación del nivel de la señal acústica en decibeles (eje *y*) para cada componente de frecuencia en hertz (eje *x*), indicando los armónicos, la envolvente espectral y las formantes.

E. Área de resultados. Contiene los resultados de: 'Formante cantante', 'Sintonía formante-armónico', 'Vibrato' y 'Entonación'.

Se indica si existe formante del cantante por medio de un gráfico.

En el área de sintonía formante – armónico, se indica el porcentaje de desajuste del formante correspondiente (F1 o F2) con el armónico más cercano. Se representa en una barra medidora, que va de -50% a +50%, pasando por cero que es el valor óptimo.

En el área de vibrato se muestra en una barra medidora, que va desde 0 hasta 12 hertz, la frecuencia del vibrato obtenida.

En el área de entonación, se muestra el nombre de la nota musical correspondiente a la frecuencia fundamental F0 del espectro, en relación con la nota *La* de 440 Hz. Se indica el ajuste o desajuste en cents, en una barra medidora que va de -50 a +50 cents.

F. Teclado de rango vocal musical. En este gráfico se indica la nota musical correspondiente.

G. Barra de estado.

⁸ El diseño correspondiente a la ventana de Ayuda, se muestra en el apéndice C de esta tesis.




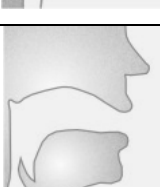
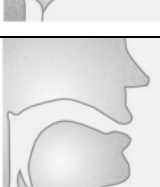
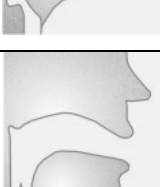
Valor de F1	Valor de F2	Posición de la lengua horizontal/vertical	Apertura bucal	Imagen
600 - 900	1000 - 2500	Central / baja	Muy abierta	
400 - 600	1500 - 2500	Anterior / media	Semiabierta	
200 - 400	1500 - 3000	Anterior / alta	Casi cerrada	
400 - 600	500 - 1500	Posterior / media	Semiabierta	
200 - 400	500 - 1500	Posterior / alta	Casi cerrada ó Cerrada	
< 200 ó >900	< 500 ó >3000	Plana	Abierta	

Tabla 2.3 Esquemas de la forma o configuración del tracto vocal, de acuerdo a la relación vocálica F1-F2.

CAPÍTULO 3. USO Y EVALUACIÓN DEL PROGRAMA “EVOCANTO”

Ya terminado el desarrollo del programa *Evocanto*, se realizaron experimentos para comprobar su funcionamiento con distintos usuarios. En este capítulo se muestran algunos de los resultados obtenidos. También, se proponen algunas formas de utilizar *Evocanto*, con detalles técnicos y con actividades para el entrenamiento de la técnica vocal. Por último, se presentan los resultados de la evaluación de *Evocanto*, hecha por parte de usuarios, a quienes se les aplicó un cuestionario de opinión, después de haber utilizado el programa.

3.1 Guías de uso y sugerencias didácticas

Para utilizar *Evocanto*, es necesario realizar correctamente la configuración de la entrada de audio a la computadora a través de un micrófono. Si se utiliza una interfaz de audio o tarjeta de sonido externa a la computadora, se debe conectar y configurar de acuerdo a las especificaciones técnicas de dicho dispositivo. Posteriormente, se necesita indicar a la computadora que tome la línea de micrófono de dicha interfaz como la entrada de audio predeterminada. En sistemas operativos Windows esto se indica en las propiedades de grabación, en el control general de volumen, al seleccionar el *dispositivo mezclador* (interfaz de audio correspondiente) y habilitar la opción de *micrófono*.

Los controles de la pantalla de *Evocanto* están diseñados para mostrar los resultados del análisis vocal. El usuario puede cambiar estos resultados solamente con su voz al micrófono. Las funciones que puede ejecutar el usuario al dar “click” con el *mouse* de la computadora en determinadas partes del programa, son: *Iniciar o Detener el análisis* – con el botón de encendido y apagado; *Limpiar pantalla* – doble click sobre el área de imagen del tracto vocal; *Cerrar el programa* – click en el botón superior izquierdo; y visualizar la ventana de ayuda – click en el botón de ayuda “?”.

Se recomienda que, para practicar con *Evocanto*, se realicen con anticipación ejercicios de relajación muscular y se adopte la postura corporal correcta para comenzar a cantar. Incluso se pueden realizar algunos ejercicios de respiración y/o vocalizaciones, para dar cierta flexibilidad a la voz; empezar con ejercicios de menor dificultad o con tonos medios, tratando de nunca forzar el mecanismo de la voz en la emisión. Se pueden realizar dichos ejercicios de acuerdo a las recomendaciones del profesor de canto o algún

método existente (existen ejercicios preparatorios en [Roe, 1970], [Collins, 1999] y [Ferrer, 2008]).

A continuación se enumeran algunas actividades que se pueden realizar con *Evocanto*, como complemento del entrenamiento vocal.

1. Cantar las cinco vocales en los diferentes registros de la voz propios de cada cantante. Localizar el punto de cambio entre dichos registros, es decir, las “notas de paso” y verificar los cambios que se suscitan en la forma o configuración del tracto vocal.
2. Localizar las vocales y alturas con mayor dificultad técnica. Comparar los cambios del tracto vocal y del espectro al realizar cambios en la apertura bucal, en el velo del paladar, en el descenso de la mandíbula, en la contracción o distensión de la lengua y en la protrusión y movimiento de los labios.
3. Entonar vocales e incluso consonantes sonoras (/m/, /n/). Observar los cambios que se producen en el espectro. Notar los cambios de energía y variaciones en la entonación al iniciar y finalizar una nota. Sostener sonidos por tiempos variables de duración, desde 1 segundo a varios segundos⁹.
4. Cantar con distinto nivel de sonoridad, es decir, aumentar y disminuir la intensidad o volumen de la voz. En voces graves, verificar la aparición de formante del cantante y en voces agudas verificar la sintonía de las formantes con los armónicos, al cantar con un alto nivel de energía. Experimentar cambios de sonoridad y compararlos con cambios en la configuración del tracto vocal.
5. Comprobar el nivel de vibrato. Distinguir que conforme se entrena la voz de modo correcto, el vibrato oscilará en niveles “normales” (ni poco, ni mucho), es decir, entre 5 y 7 Hz, en el centro de la barra deslizante indicadora del vibrato. Hacer cambios voluntarios en el nivel de vibrato y verificar los cambios en pantalla. *Evocanto* sólo registrará cambios en la frecuencia de vibrato si el vibrato se realiza de forma natural en la fuente glotal de voz con apoyo en el diafragma.
6. Practicar la entonación de las notas musicales y su ajuste. Hacer que el ajuste en cents de la nota musical quede cercano a cero (0 cents), en la barra medidora, ya que esto indica una entonación exacta. Aún ajustes de entre unos 5 cents positivos o negativos pueden considerarse normalmente como excelentes.

⁹ Evocanto tiene un tiempo de análisis mínimo de 0.7 s (siete décimas de segundo), que no permite analizar tonos muy cortos.

7. Realizar pruebas de entonación con diferentes intervalos musicales. Verificar los cambios que se producen al pasar de un tono a otro; hacer tonos ligados y picados (con duración mínima de 1 segundo).

Todas estas actividades son sólo algunas sugerencias. Sin embargo, se pueden crear otras u otra manera de aplicarlas, de acuerdo al contexto en que se utilice el programa (personal, académico, educativo, etc.), así como al criterio del profesor o a la práctica y avance del alumno de canto.

3.2 Pruebas técnicas y ejemplos de aplicación

Para implementar *Evocanto*, se utilizaron los siguientes elementos:

- una computadora tipo laptop, marca Toshiba, de procesador Intel Celeron, con sistema operativo Windows XP, en la cual se instaló *Evocanto*;
- una interfaz de audio marca Edirol, modelo UA-25, configurada en modo de compatibilidad: 16 bits, 44.1 kHz, conectada a la computadora a través del puerto USB;
- un micrófono direccional dinámico super-cardioide AKG Acoustics modelo D880M (especial para canto) con conector XLR de tres polos, conectado a una entrada balanceada de la interfaz de audio, con un cable XLR balanceado.

Una vez conectado, encendido y configurado el sistema de *Evocanto*, se comprobó su funcionamiento y se realizaron algunas pruebas, por parte de diferentes cantantes, representativos de cada tipo de voz.

Para llevar a cabo las pruebas, se le pidió a cada cantante entonar al micrófono las cinco vocales con una nota del registro bajo, una del registro medio y una del registro agudo, sosteniendo cada vocal por aproximadamente siete u ocho segundos. Posteriormente, se les pidió entonar algunas notas de distintos registros y con diferentes vocales, tratando de disminuir el vibrato y después con vibrato exagerado. Fueron sugeridas las siguientes notas para cada tipo de voz:

Voz/Registros	bajo	medio	agudo	Voz/Registros	bajo	medio	agudo
Soprano	Re#5	La#5	Fa6	Tenor	Re#4	La#5	Fa6
Mezzosoprano	Do5	Sol5	Re6	Barítono	Do4	Sol4	Re5
Contralto	La4	Mi5	Si5	Bajo	La3	Mi4	Si4

A continuación se ilustran algunos de los resultados obtenidos, de las voces Mezzo-soprano y Barítono¹⁰.

¹⁰ Se eligieron éstas por ser representativas de la clasificación vocal intermedia de mujer y hombre, respectivamente, y porque son contrastantes en sus características tímbricas.

Mezzo-soprano

Las siguientes figuras corresponden a las capturas de pantalla de *Evocanto* en funcionamiento, siendo utilizado por una cantante mezzo-soprano de alto nivel. En la figura 3.1 se observa la vocal /o/ cantada en los registros grave, medio y agudo. Las envolventes espectrales se muestran equivalentes, con la correspondiente separación de armónicos entre sí, debido al aumento de frecuencia. La amplitud o nivel de energía se conserva de modo estable. Es notoria la elongación del tracto vocal, mediante la protrusión de los labios, como puede observarse en los esquemas de configuración del tracto vocal. También existe sintonía de la primera formante F1 con el armónico más cercano. El vibrato aumenta en el registro agudo; y en el registro grave existe formante del cantante.

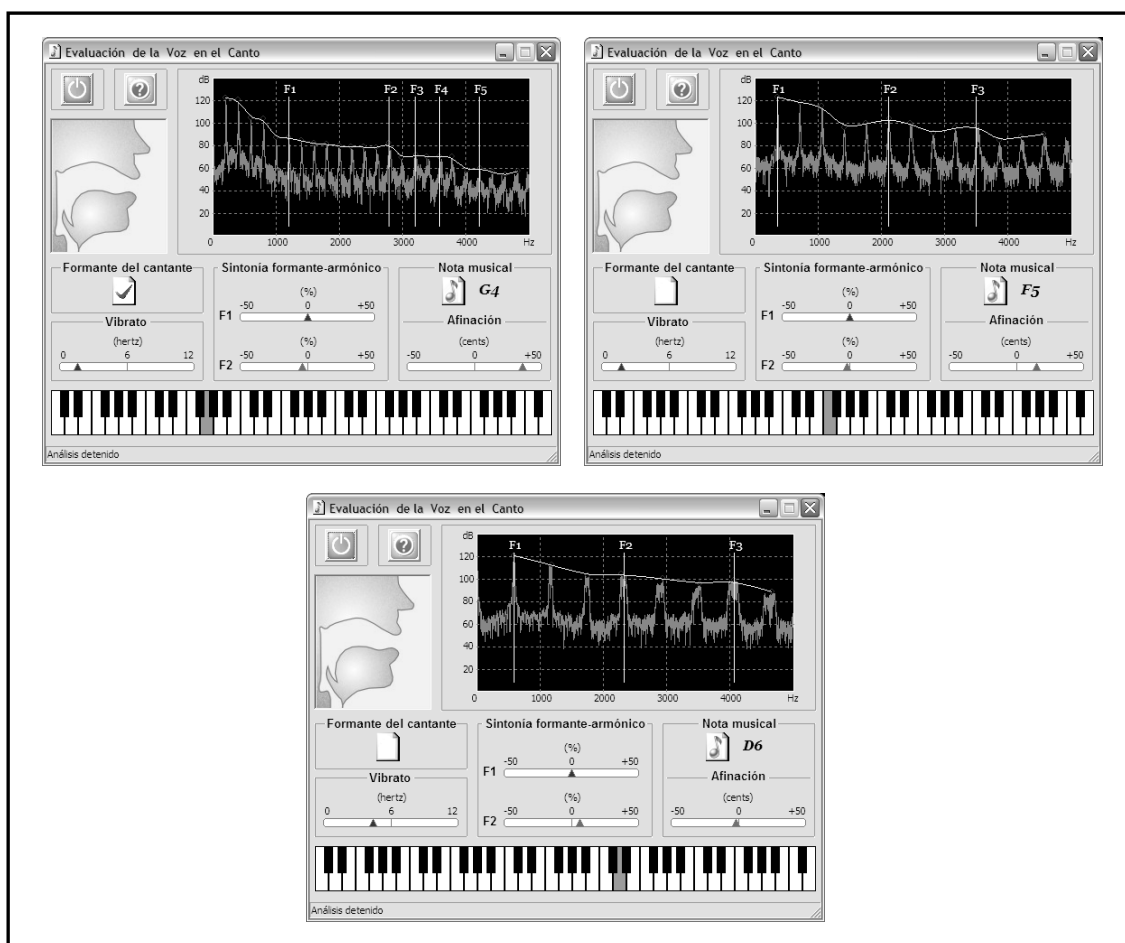


Figura 3.1 Vocal /o/ en los registros grave, medio y agudo en voz de mezzosoprano..

En la figura 3.2 son presentadas las cinco vocales cantadas en el registro medio de la voz de mezzosoprano. No hay cambios notables en el nivel de energía. Permanece la tendencia a sintonizar las formantes con armónicos. Las formas de las envolventes espectrales son muy características, puede observarse un gran parecido entre la vocal /i/ y la vocal /u/, así como entre la vocal /a/ y la /o/. Esto mismo se observa en las configuraciones del tracto vocal.

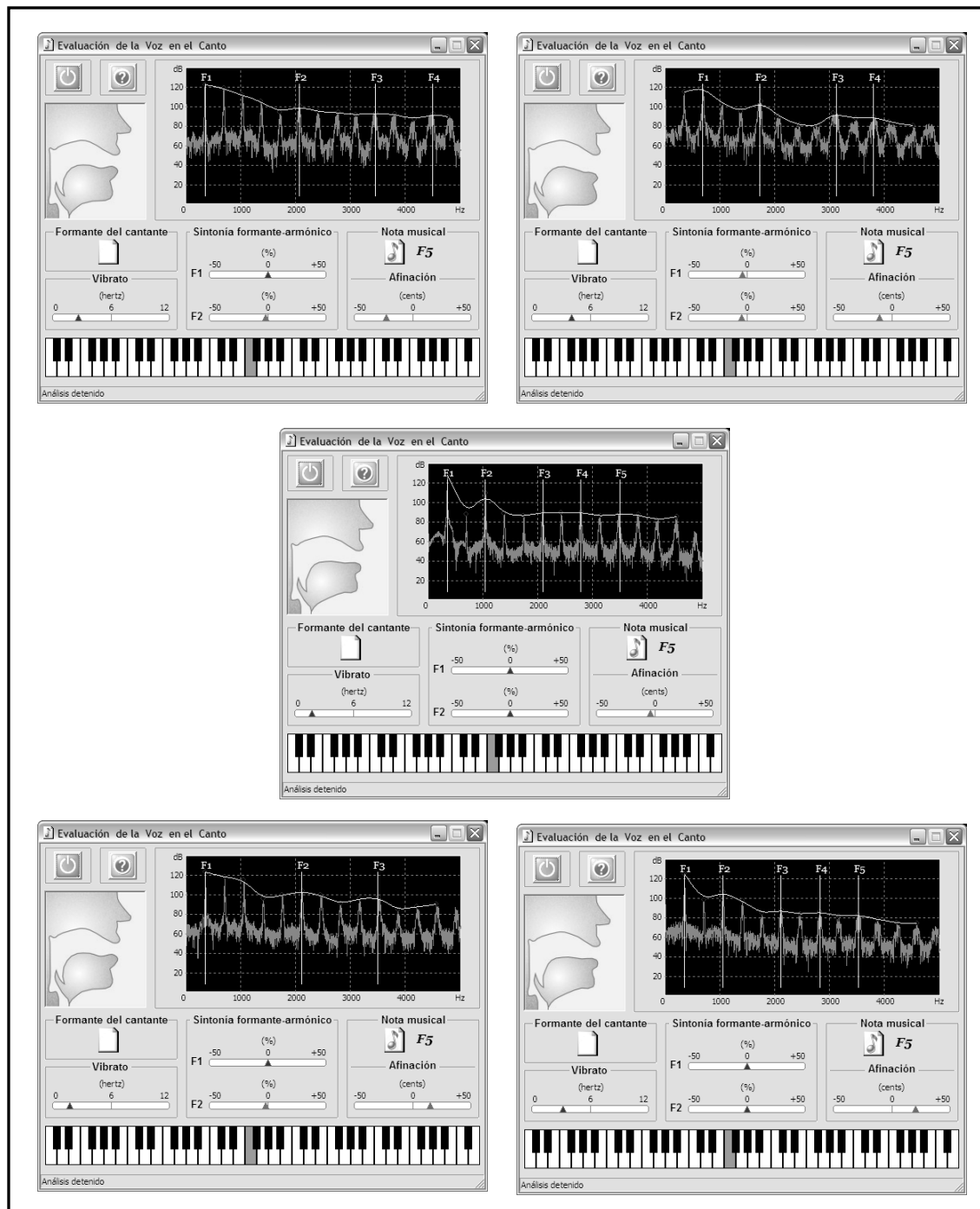


Figura 3.2 Vocales /a/,/e/,/i/,/o/,/u/ en el registro medio de mezzosoprano.

De acuerdo a lo observado, la vocal /e/ presenta un cambio en la configuración del tracto vocal de la cantante mezzo-soprano. Al parecer existe una mayor apertura bucal y/o una posición centrada horizontalmente de la lengua. Esto se reafirma al observar la captura de pantalla de la misma vocal /e/ en el registro agudo, donde se aprecia la sintonía de la formante F1 con el armónico alto, lo cual indica la posición hacia abajo de la mandíbula. En la figura 3.3 se observan las semejanzas en los registros medio y agudo de la vocal /e/.

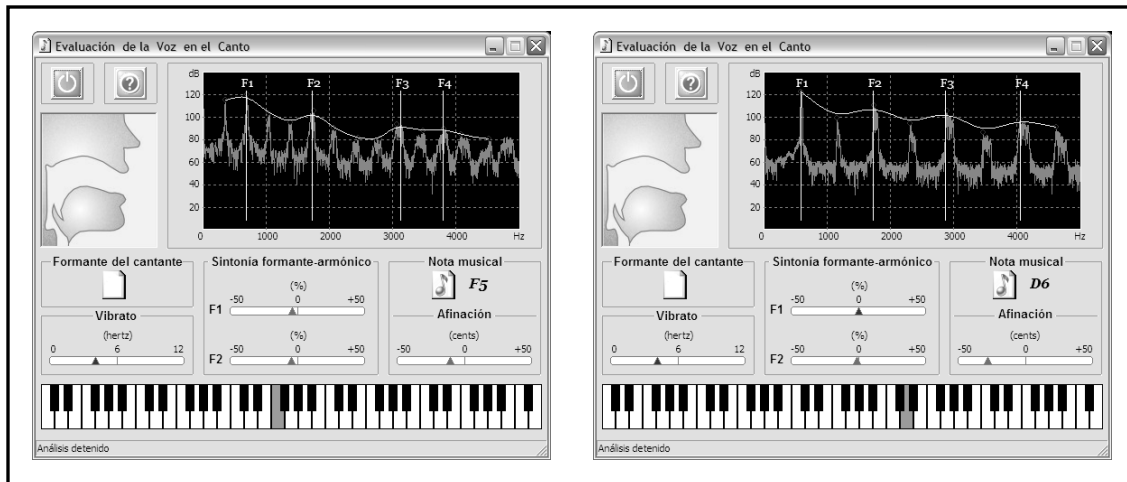


Figura 3.3 Vocal /e/ en los registros medio y agudo de mezzosoprano.

Por último, en la figura 3.4 se muestran las diferencias en la medida del vibrato, realizadas en forma voluntaria, por la misma mezzo-soprano. Para lograr disminuir o aumentar el vibrato, la cantante realizó cambios en la forma o configuración de su tracto vocal, afectando también la amplitud en modo proporcional al nivel de vibrato.

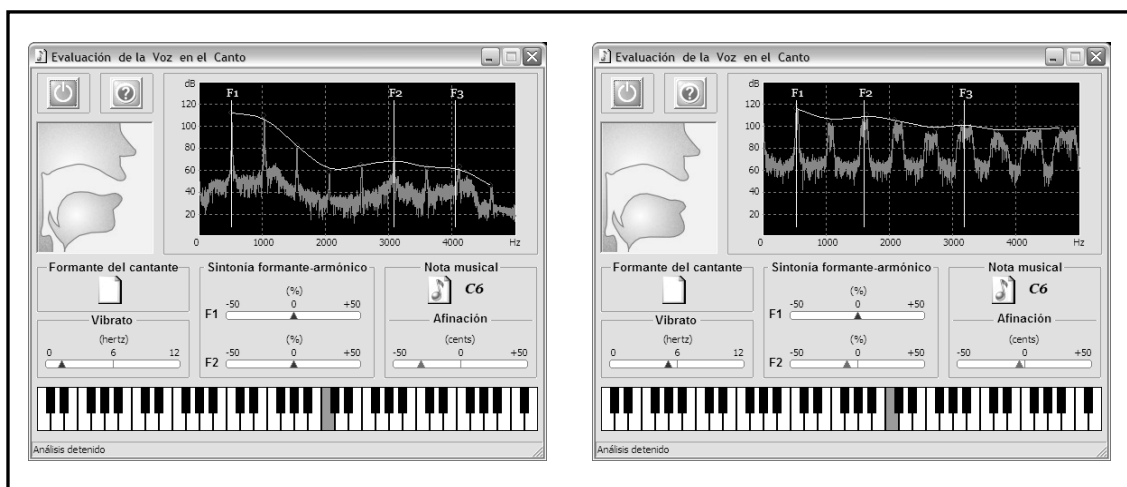


Figura 3.4 Vocal /a/ con poco y mucho vibrato de mezzosoprano.

Barítono

En la figura 3.5 se muestran las capturas de pantalla de *Evocanto*, que corresponden a la vocal /a/, cantada en los registros grave, medio y agudo de un cantante barítono profesional. Las diferencias determinadas en los espectros, muestran cambios en la amplitud; sin embargo la configuración del tracto vocal es semejante en los tres casos, por tratarse de la misma vocal. Se observa que, conforme se dirige la voz al registro agudo, se va extendiendo la distribución de los armónicos, es decir, se presentan más espaciados entre sí.

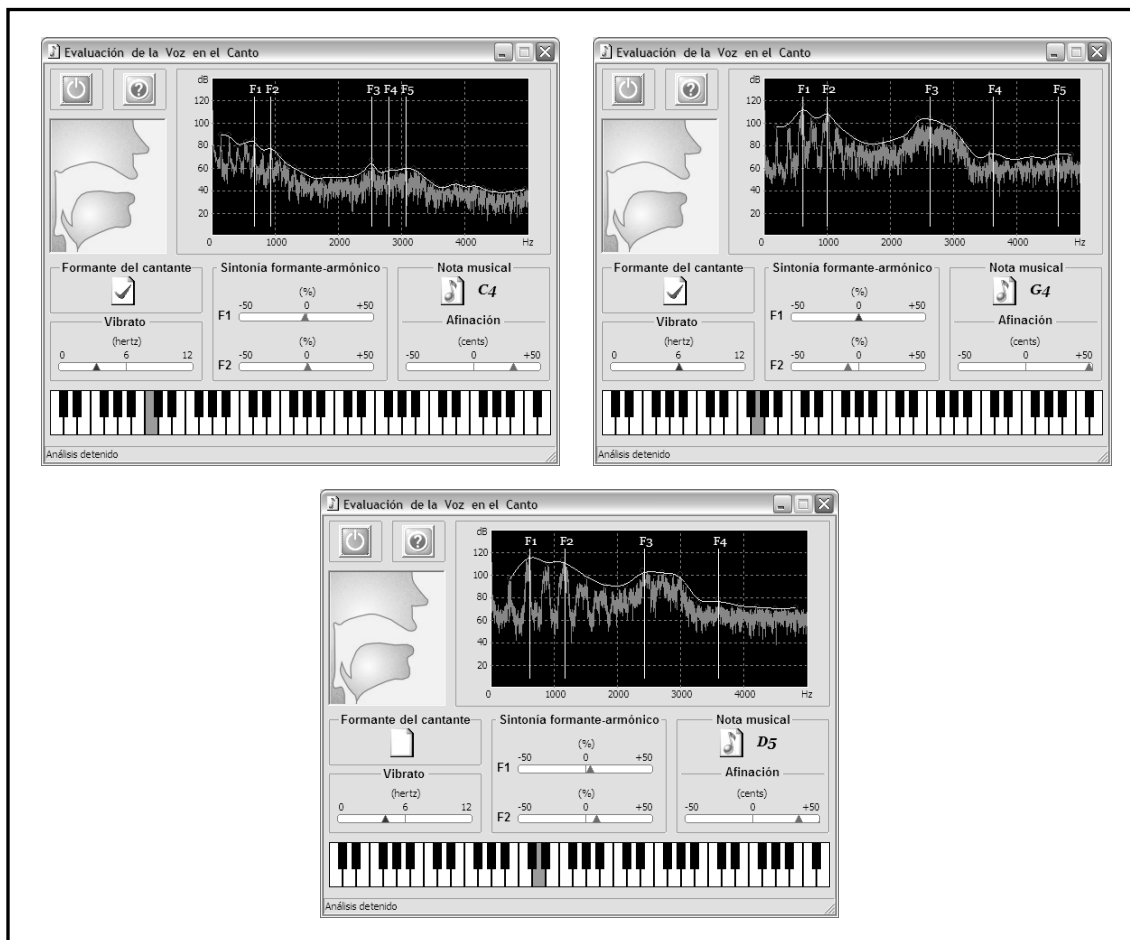


Figura 3.5 Vocal /a/ cantada en los registros grave, medio y agudo de voz barítono.

La figura 3.6 muestra capturas de pantalla de *Evocanto* al analizar las cinco vocales, cantadas por el barítono. Se puede apreciar la diferencia en la envolvente espectral de cada vocal; en la distribución, magnitud y relación de las formantes, principalmente F1 y F2, que son las que determinan la configuración del tracto vocal como se muestra en los esquemas correspondientes; y en las siguientes formantes, a partir de

F3, se distingue la aparición de la formante del cantante. También presenta una medida de vibrato intermedia, propia de cantantes profesionales.

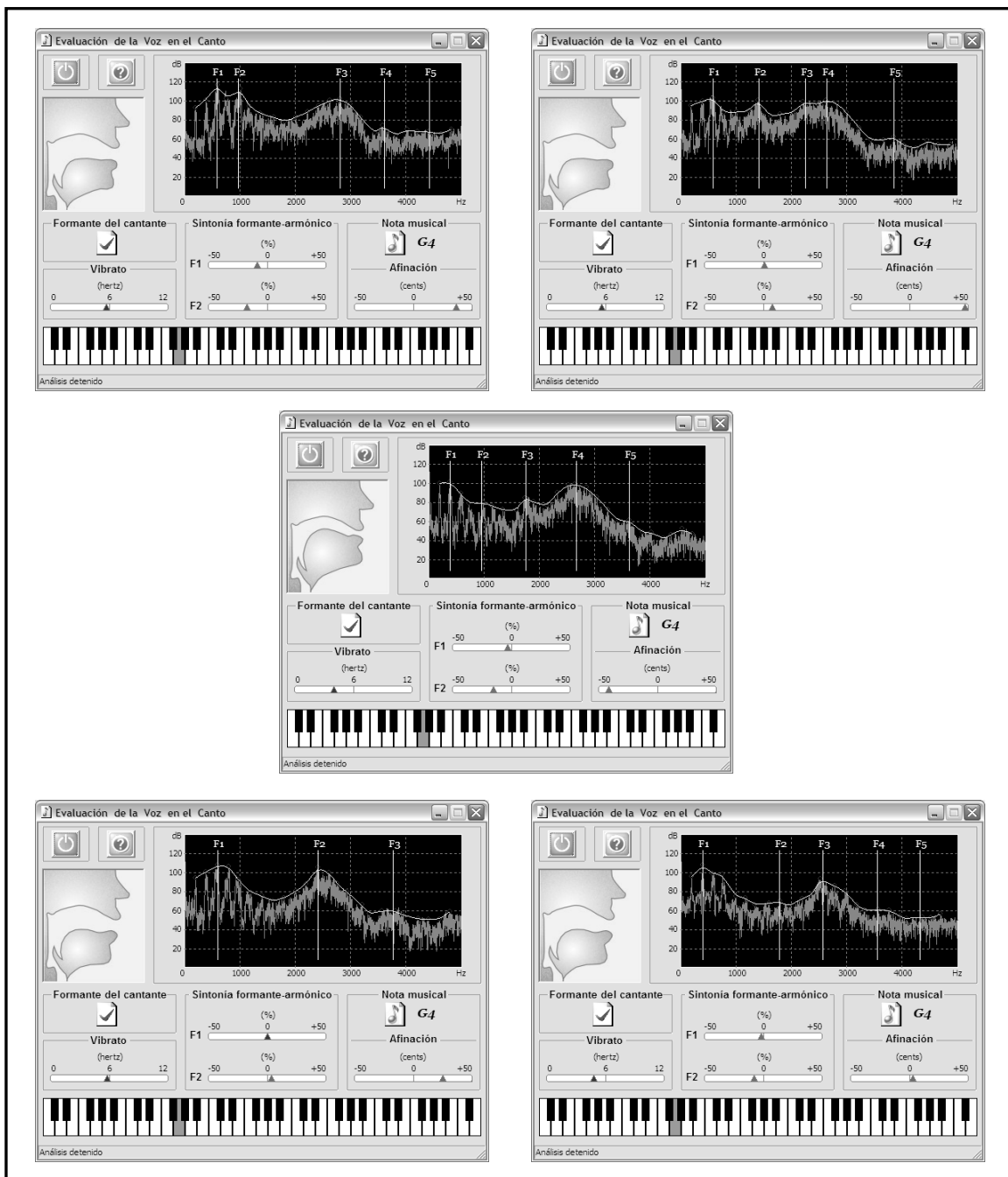


Figura 3.6 Vocales /a/,/e/,/i/,/o/,/u/ en el registro medio de la voz de barítono.

La figura 3.7 presenta las capturas de pantalla de *Evocanto*, con la diferencia en la cantidad de vibrato de la vocal /o/, cantadas por el mismo barítono, en una nota del registro medio de la voz. Se aprecian ligeros cambios en el espectro, debido a esta diferencia. Existe un buen control del vibrato a voluntad, por parte del cantante, pues las configuraciones del tracto vocal se conservan.

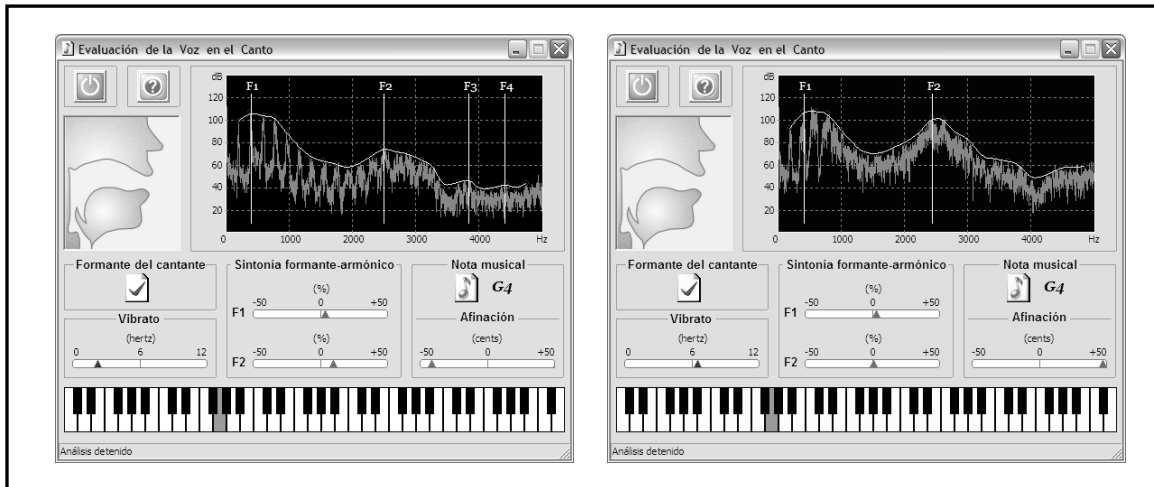


Figura 3.7 Diferencias en poco y mucho vibrato de la voz barítono.

3.3 Pruebas de opinión

Al finalizar una breve sesión de capacitación y uso del programa, se aplicó un cuestionario para obtener la opinión de los usuarios de *Evocanto*, donde se evaluaron los principales criterios de uso y desempeño del programa, de acuerdo a las siguientes características: interfaz gráfica, información, facilidad operacional, tiempo de respuesta, claridad y uso didáctico.

Dichas características se trataron en diez aspectos del programa (conformando diez preguntas o aspectos a evaluar), que el usuario debía calificar, con base en la siguiente escala de ponderación, siendo 4 la mayor calificación y 0 la más baja.

- Excelente = 4 puntos
- Buena = 3 puntos
- Suficiente = 2 puntos
- Pobre = 1 punto
- Mala = 0 puntos

A continuación se presenta un análisis general de los resultados obtenidos en cada elemento del cuestionario, aplicado a distintos cantantes profesionales: estudiantes (principiantes y avanzados) y profesores.

1. **Distribución de los elementos de la interfaz gráfica.** Se trata de la apariencia estética del programa. Los usuarios calificaron este aspecto, de acuerdo a cuál fue su primera impresión al visualizar la pantalla. A la mayoría de los usuarios les pareció muy bien en tamaño y distribución, sólo hubo contadas excepciones que preferirían que el botón de encendido del programa o el teclado tuviesen una distribución diferente.
2. **Información que ofrece acerca de la voz cantada.** Para algunos usuarios, los conceptos y parámetros que presenta el programa son un tanto difíciles de comprender después de sólo una breve sesión de capacitación. Sin embargo, coinciden en que después de haber aprendido a utilizar el programa, fue más fácil entender de qué se trataban los conceptos técnicos de la voz.
3. **Facilidad de uso.** Hubo algunos usuarios que dudaron en la facilidad de uso, debido a su falta de conocimiento acerca de los conceptos de acústica de la voz.
4. **Tiempo de respuesta del análisis.** Algunos usuarios no están completamente satisfechos con la velocidad en la presentación de los resultados del análisis vocal; para unos es lento, para otros es rápido.
5. **Claridad de los conceptos relativos a la voz cantada.** La mayoría de usuarios que conocen algo acerca de los conceptos tratados, opinaron que existe un buen nivel de claridad. Para los que los conceptos son desconocidos, coinciden en que necesitan familiarizarse con ellos.
6. **Evaluación como herramienta para el estudio del canto.** Los usuarios estudiantes coincidieron en que sería un buen complemento para su entrenamiento, pero sabiendo utilizar integralmente el programa para “sacarle provecho”. A la mayoría de los profesores consultados les agradó mucho el programa, y expresaron interés en poder aplicarlo en sus clases.
7. **Como herramienta para mejorar las habilidades vocales.** Se coincide en que es una buena herramienta para practicar las características vocales de las que trata el programa, sin embargo, no es útil para otras habilidades vocales-musicales, por ejemplo las que requieren velocidad en la ejecución.

8. **Como herramienta para mejorar las habilidades auditivas.** La mayoría de usuarios considera que para poder lograr una ayuda en la mejora de las habilidades auditivas, el programa debería tener la opción de grabar la voz cantada.
9. **Como herramienta didáctica para enseñanza en grupo.** Los usuarios opinaron que sería divertido experimentar qué pasa con el funcionamiento del programa si fuese utilizado por dos o más voces en conjunto. Pero, que es muy obvio que el programa está diseñado para ser utilizado en forma individual.
10. **Como ayuda para estudiar en forma independiente.** La opinión de los usuarios es determinante en que se necesita una guía didáctica; serían capaces de utilizar el programa en forma independiente siempre y cuando tengan en claro los objetivos perseguidos y la forma de llegar a ellos; otros opinan sólo se requiere una mayor capacitación y entrenamiento en el uso del programa.

La estadística de los resultados obtenidos arrojó que en opinión de los usuarios, el desempeño del programa en general, es **bueno**, ya que se obtuvo un promedio de 3.4 de calificación, con un 85% de nivel medio de aceptación. En la siguiente tabla se observan los promedios por cada aspecto del cuestionario (Tabla 3.3).

CRITERIO	CALIF. PROMEDIO	PORCENTAJE DE ACEPTACIÓN
Distribución de los elementos de la interfaz gráfica	3.75	94%
Información que ofrece acerca de la voz cantada	3.75	94%
Facilidad de uso	3.75	94%
Tiempo de respuesta del análisis	3.25	81%
Claridad de los conceptos relativos a la voz cantada	3.37	84%
Evaluación como herramienta para el estudio del canto	3.75	94%
Como herramienta para mejorar las habilidades vocales	3.37	84%
Como herramienta para mejorar las habilidades auditivas	2.87	72%
Como herramienta didáctica para enseñanza en grupo	2.62	66%
Como ayuda para estudiar en forma independiente	3.62	90%

Tabla 3.3 Resultados de las pruebas de opinión, acerca del desempeño de Evocanto.

CONCLUSIONES

Aportaciones y alcances

Una de las aportaciones más importantes de este trabajo, es la de proponer y materializar la idea de elaborar un programa de cómputo que analiza la voz cantada y que está dirigido a cantantes profesionales (alumnos y profesores) como una herramienta de trabajo, para ayudar en la práctica y desarrollo de la técnica vocal. Mediante la utilización de *Evocanto*, el usuario contará con un complemento de su formación como cantante, que contribuirá al logro de objetivos técnicos.

A través del lenguaje de programación Python y las bibliotecas SndObj y WxWidgets, de código libre, se generó *Evocanto* como un programa multiplataforma, de interfaz gráfica amigable hacia el usuario. Mediante este programa el usuario puede observar lo que ocurre con su voz en el momento mismo de cantar, lo que permite orientarlo hacia un adecuado balance entre la funcionalidad de las estructuras anatómicas, particularmente del sistema vocal, y los aspectos estético-perceptivos implicados en los parámetros acústicos de la voz.

El análisis acústico de la voz cantada, que se realiza en el programa *Evocanto*, consiste en la obtención del espectro de la señal de voz para detectar elementos característicos, como son: armónicos, envolvente espectral, la frecuencia fundamental F0 y las frecuencias formantes F1, F2,.. y la amplitud. Con estos elementos se logran indicar los siguientes objetivos de evaluación del análisis vocal:

- Diagrama de la forma o configuración del tracto vocal.
- Verificación de la existencia de formante del cantante y del grado de sintonización de formantes con armónicos.
- Evaluación de entonación y vibrato, proporcionando la información correspondiente a la nota musical cantada y a la cantidad o frecuencia de vibrato.

Para lograr que *Evocanto* realice sus funciones con eficiencia y la mayor exactitud posible, fueron implementados algoritmos particulares para el análisis del espectro en la obtención de la frecuencia tonal F0 y frecuencias formantes F1, F2, ..., para el cálculo de la nota musical y para el vibrato. La creación e implementación de estos algoritmos es una aportación original de este trabajo.

Asimismo, la forma en que se presentan los resultados en la pantalla de *Evocanto*, es una aportación original, que se realizó con la finalidad de mantener una interfaz gráfica sencilla, coherente y funcional, donde se presentan esquemas del tracto vocal, el esquema del espectro de voz con todos sus elementos y los indicadores de ajuste con un elemento gráfico común, que es una barra con aguja de medición, cuyo punto de desempeño óptimo se ubica en el centro de la barra.

Los usos didácticos (sección 3.2) propuestos para la implementación de *Evocanto*, son sólo sugerencias en consideración de la funcionalidad del programa, pues no es tema de esta tesis el aspecto didáctico, que puede ser profundizado en otros estudios.

Se evaluó *Evocanto* con distintos tipos de voz y usuarios, aplicando pruebas de opinión, que dieron un resultado aceptable y satisfactorio acerca del desempeño y de la interfaz gráfica del programa.

Sugerencias de trabajo a futuro

Las ideas que pueden llevarse a cabo en trabajos posteriores, basados en *Evocanto*, o como continuación de este trabajo, son:

- Implementar la función de grabación de archivos de voz; para la percepción auditiva del propio usuario y proveer un historial de avances.
- Utilizar *Evocanto* en un curso escolar de canto profesional o técnica vocal.
- Integración de ejercicios vocales de acuerdo a un método progresivo.
- Adaptar la interfaz gráfica del programa *Evocanto* de acuerdo al nivel de los usuarios cantantes (principiantes, intermedios, avanzados) y/o a las clasificaciones (tesituras) de voz.
- Mejorar los esquemas de configuración del tracto vocal.
- Perfeccionar los algoritmos de la detección de frecuencia tonal y de formantes.
- Determinación de la vocal que se está cantando.
- Crear e incorporar nuevos algoritmos para la medición de otros aspectos del vibrato y del timbre de la voz cantada.

REFERENCIAS

- Benade, Arthur. *Fundamentals of Musical Acoustics*. Chap. 19: "The Voice as a Musical Instrument". Second revised edition. Dover, 1990.
- Blades-Zeller, Elizabeth. *Singing with Your Whole Self: The Feldenkrais Method and Voice*. The Scarecrow Press, 2002.
- Bloothoof, G; Plomp, R. *The sound level of the singer's formant in professional singing*. En: The Journal of the Acoustical Society of America, Vol. 79, no. 10 (pp. 2028 – 2033). 1986.
- Bunch, Meribeth. *Dynamics of the singing voice*. Springer. Viena, 1997.
- Capponi, A. *La evaluación en el Canto: categorías de estimación y habilidades de ejecución vocal* Universidad Nacional de La Plata, Buenos Aires, Argentina, 2004.
<http://www.sacom.org.ar/secciones/cuarta/resumenes.htm>, fecha de consulta: 17 de diciembre de 2008.
- Chowning, John. *Computer Synthesis of the Singing Voice. Research Aspects on Singing*. KTH, Estocolmo, Suecia, 1981.
- Canuyt, G., *La voz*. Hachette. Buenos Aires, 1958.
- Collins, Don L. *Teaching Choral Music*. Prentice Hall, 1999.
- Cordourier, Héctor; Orduña, Felipe. *Active control of periodic fan noise in laptops: spectral width requirements in a delayed buffer implementation*. En: Journal of applied research and technology. Vol. 7. No. 2. August 2009. CCADET, UNAM. México.
- Cornut, Guy. *La voz*. Fondo de Cultura Económica. México, 1985.
- Dejonckere, P. *Caruso's vibrato: An acoustic study*. En: P. Dejonckere, M. Hirano, & J. Sundberg, (Eds.), *Vibrato* (pp. 111-120). Singular Publishing Group. San Diego, 1995.
- Dromey, Christopher; Reese, Lorie; Hopkin, J. Arden. *Laryngeal-Level Amplitude Modulation in Vibrato*. En Journal of Voice, Vol. 23, issue 2, marzo. pp 156-163. 2009.
- Fant, Gunnar. *Acoustical Analysis of speech*. En: CROCKER, Malcom (editor). *Enciclopedia of Acoustics*. John Wiley, 1997.
- Ferrer, Joan S. *Teoría, anatomía y práctica del canto*. Herder. España, 2008.
- Gándara, Manuel *La interfaz en el usuario: una introducción para educadores* En: Usos educativos de la computadora. CISE-UNAM. México, 1994.
- García-Tapia, R. *Antecedentes Históricos*. En: *Diagnóstico y tratamiento de los trastornos de la voz*. Editorial GARSÍ. Madrid, 1996.
- Heirich, Jane. *Voice and the Alexander Technique*. Mornum Time Press, 2005.

Horii, Y.; Hata, K. *A note on phase relationships between frequency and amplitude modulations in vocal vibrato*. En: *Folia Phoniatrica*, 40, pp 303-311. 1988.

Ladefoged, Peter. *A note on "Information conveyed by vowels"*. En: *The Journal of the Acoustical Society of America*, Vol. 85, no. 5 (pp. 2223 – 2224). May 1989.

Lazzarini, Victor. *The Sound Object Library*. Cambridge University Press. Cambridge, 2000.

Lemmetty, S. *History and Development of Speech Synthesis*, en http://www.acoustics.hut.fi/publications/files/theses/lemmetty_mst/contents.html
Fecha de consulta: 08 de abril de 2008.

Lindblom, Björn; Sundberg, Johan. *The human voice in speech and singing*. Cap. 16 en Rossing, Thomas D. (editor) *Springer handbook of acoustics*. Springer, 2007.

Miller D. G.; Schutte H. K.; Sulter A. M.; Wolf R.F. *Comparison of Vocal Tract Formants in Singing and Nonperiodic Phonation*. *Journal of Voice*, Vol. 11, No. 1, pp 1-11. 1997.

Nearey, Terrance M. *Static, dynamic, and relational properties in vowel perception*. En: *The Journal of the Acoustical Society of America*, Vol. 85, no. 5 (pp. 2088 – 2113) May 1989.

Owens, Frank J. *Signal processing of speech*. McGraw-Hill, México, 1993.

Potter, R; Steinberg, J. *Toward the Specification of Speech*. En: *The Journal of the Acoustical Society of America*, Vol. 22, no. 6. Noviembre. Bell Telephone Laboratories. Murray Hill, New Jersey. 1950.

Quilis, A.; Esgueva, M. *Realización de los fonemas vocálicos españoles en posición fonética normal*. En: Esgueva y Cantarero (Eds.), *Estudios de fonética I*. (pp. 137-252). Consejo Superior de Investigaciones Científicas. Madrid, 1983.

Rabiner, Lawrence; Schafer, Ronald. *Digital processing of speech signals* Prentice Hall. New Jersey, 1978.

Robison, C.; Bounous, B.; Bailey, R. *Vocal beauty: A study proposing its acoustical definition and relevant causes in classical baritones and female belt singers*. En: *The National Association of Teachers of Singing Journal*, 42. pp 19-30. 1994.

Roe, Paul F. *Choral Music Education*. Prentice Hall, 1970.

Roederer, Juan G. *Physics and Psychophysics of Music: An Introduction*. Springer-Verlag, 1995.

Seashore, C. *The vibrato*. Iowa City: University of Iowa, 1932.

Seidner, Wolfram; Wendler, Jürgen. *Die Sängerstimme. Phoniatische Grundlagen des Gesangs*. Henschel Verlag. Berlin, 2004.

Smith, Winthrop. W. *Handbook of real-time fast Fourier transforms: algorithms to product testing*. IEEE. New York, 1995.

Stanley, Douglas. *The science of voice: an application of the laws of acoustics*. Fischer. New York, 1958.

Sundberg, Johan. *Articulatory interpretation of the 'singing formant'*. En: Journal of the Acoustical Society of America. 55.pp 838-844. 1974.

Sundberg, Johan. *The science of the singing voice*. Northern Illinois University. Illinois, 1987.

Sundberg, J.; Skoog, J. *Dependence of jaw opening on pitch and vowel in singers*. En: Journal of Voice 11, pp 301-306. 1997.

Sundberg, J. *Where does the sound come from?* En: A. Potter (Ed.), *The Cambridge Companion to Singing* (pp. 231-247). Cambridge: Cambridge University Press, 2000.

Sundberg, J. *Level and centre frequency of the singer's formant*. En: Journal of Voice 15, pp 176-186. 2001.

Titze, I. R. *The physics of small-amplitude oscillations of the vocal folds*. En: Journal of the Acoustical Society of America, 83, pp 1536–1552. 1988.

Titze, I.; Story, B.; Smith, M.; Long, R. *A reflex resonance model of vocal vibrato*. Journal of the Acoustical Society of America, 111, pp 2272-2282. 2002.

Titze I. R. *Nonlinear source--filter coupling in phonation: Theory*. En: Journal of the Acoustical Society of America, 123(5), pp 2733-2749. 2008.

Van Rossum, Guido; Drake, Fred. *The Python Language Reference Manual*. NetworkTheory. Bristol, 2003.

Vennard, William. *Singing: the Mechanism and the Technic*. Edición revisada por Carl Fisher. 1967.

Internet

[1] <http://www.python.org/> Fecha de consulta: 07 de octubre de 2008

[2] <http://www.riverbankcomputing.co.uk/pyqt/> Fecha de consulta: 17 de octubre de 2008

[3] <http://www.pygtk.org/> Fecha de consulta: 17 de octubre de 2008

[4] <http://www.python.org/windows/pythonwin/> Fecha de consulta: 17 de octubre de 2008

[5] <http://www.wxpython.org/> Fecha de consulta: 25 de noviembre de 2008

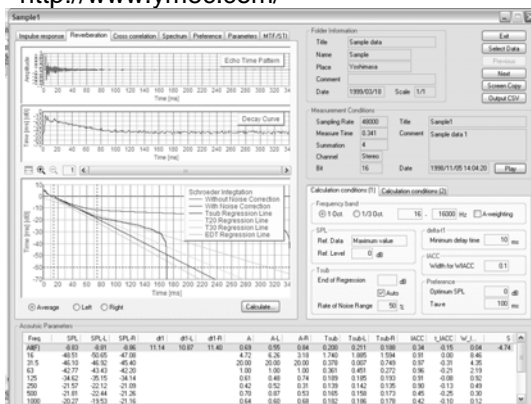
[6] <http://www.pysndobj.org/> Fecha de consulta: 25 de noviembre de 2008

APÉNDICE A

1. Programas para análisis acústico de señales de audio

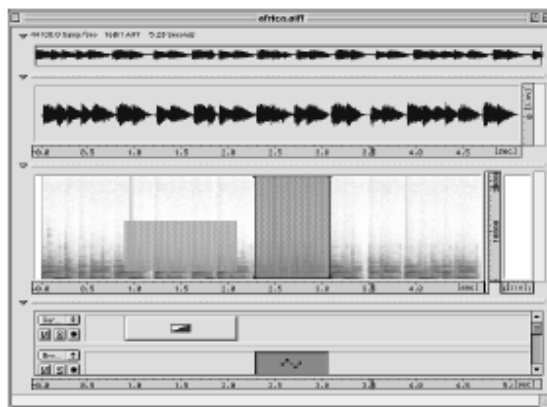
Acoustic Analyzing System 5.1

País, año: Japón, 2000
 Tipo de licencia: Comercial. El precio depende de los módulos que se deseen adquirir. Programa básico US\$990
 Campo de aplicación: Análisis de señales de audio
 Usuarios: Expertos en acústica
 Funciones: Osciloscopio, generador de señales, analizador FFT, respuesta a impulso, reverberación, ruido, espectrograma, grabación.
 Sistema operativo: Windows 95/98/ME/2000/NT4.0
 Referencia: <http://www.ymec.com/>
 Pantalla principal:



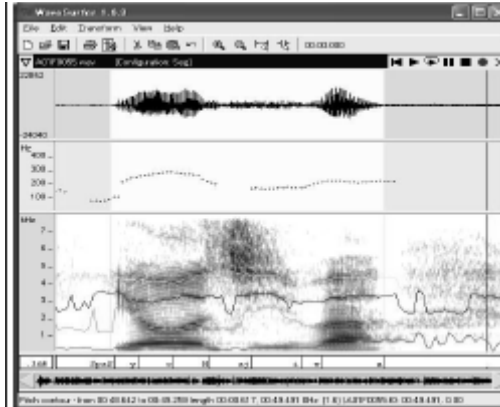
AudioSculpt 2.9

País, año: Francia, 2007
 Tipo de licencia: Se paga por suscripción, individual o institucional. Los precios van desde 39 € hasta 1100 €
 Campo de aplicación: Edición de audio
 Usuarios: Músicos expertos en manejo de audio digital
 Funciones: Manipulación de audio: filtros, cortar, copiar, síntesis, compresión. Presentación gráfica del sonido en espectrograma y sonograma (FFT, LPC, envolvente espectral y formantes)
 Sistema operativo: Mac OSX 10.4 (Tiger) o mayor. Universal Binary, Mac PPC (G4 o mayor), Mac Intel Compatible.
 Referencia: <http://forumnet.ircam.fr/349.html>
 Pantalla principal:



WaveSurfer 1.8.5

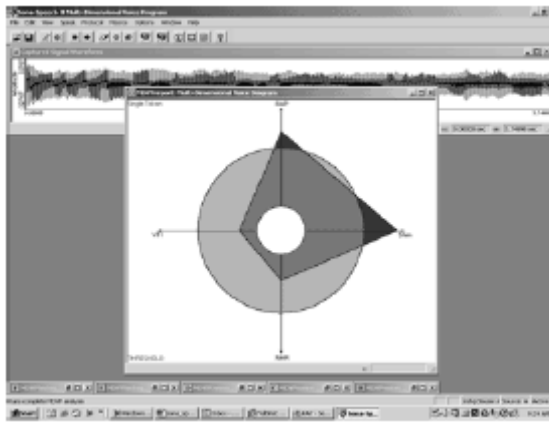
País, año: Suecia, 2005
Tipo de licencia: BSD (abierto o libre)
Campo de aplicación: Edición de audio digital
Usuarios: Profesores, investigadores.
Funciones: Espectrograma y análisis de tono. Manipulación de audio.
Sistema operativo: Linux, Windows 95/98/NT/2K/XP, Macintosh, Sun Solaris, HP-UX, FreeBSD, and SGI IRIX
Referencia: <http://www.speech.kth.se/wavesurfer>
Pantalla principal:



2. Programas para análisis médico de la voz

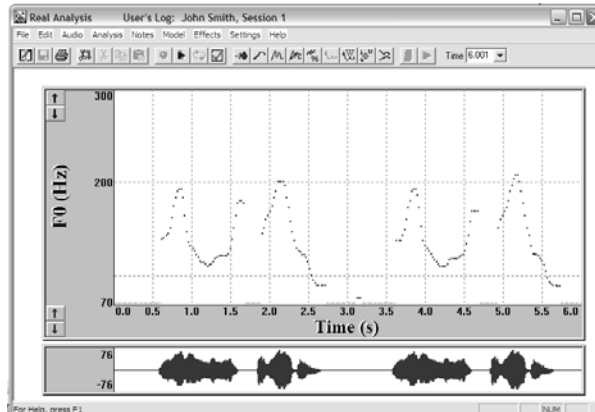
Computerized Speech Lab (CSL) Model 4500

País, año: EU, 2006
Tipo de licencia: Comercial.
Campo de aplicación: Análisis de habla
Usuarios: Profesores, científicos y especialistas en el área.
Funciones: Análisis, detección de patologías, electroglotograma, espectrograma, videos, simulaciones, bases de datos. Distribuido en diversos módulos.
Sistema operativo: Windows XP, Vista. Requiere hardware especializado.
Referencia: www.kaypentax.com
Pantalla principal:



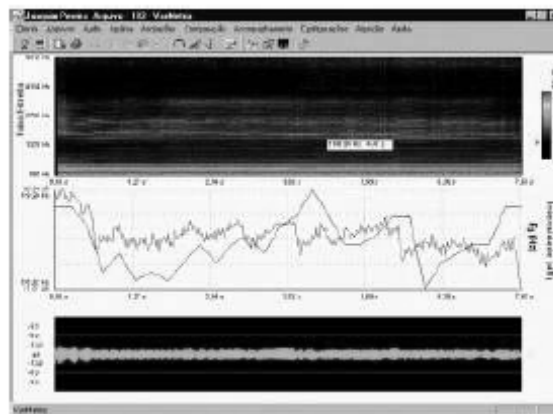
Dr. Speech 4

País, año: EU, y China, 1999
Tipo de licencia: Comercial
Campo de aplicación: Detección de patologías del habla.
Usuarios: Especialistas en el área.
Funciones: Diversos módulos: análisis en tiempo real, detección de tono, formantes, fonetograma, electroglotograma, terapias.
Sistema operativo: Windows 95/NT/98/2000/XP
Referencia: <http://www.drspeech.com/>
Pantalla principal:



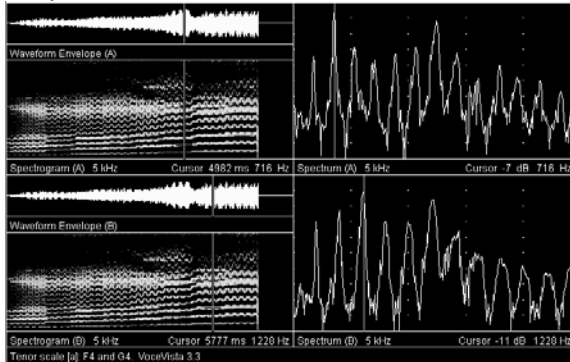
Voxmetria

País, año: Brasil, 2009
Tipo de licencia: Comercial. Precio R\$ **920,00**
Campo de aplicación: Fonoaudiología y Fisioterapia
Usuarios: Especialistas clínicos (fonoaudiólogos)
Funciones: Edición de audio, análisis de voz, histograma, espectro, formantes F1 y F2, estadísticas y funciones de calidad vocal (diagrama de desvío fonatorio, historial, comparaciones)
Sistema operativo: Windows XP/Vista/Vista 64/7
Referencia: <http://www.ctsinformatica.com.br/voxmetria.htm>
Pantalla principal:



Voce Vista

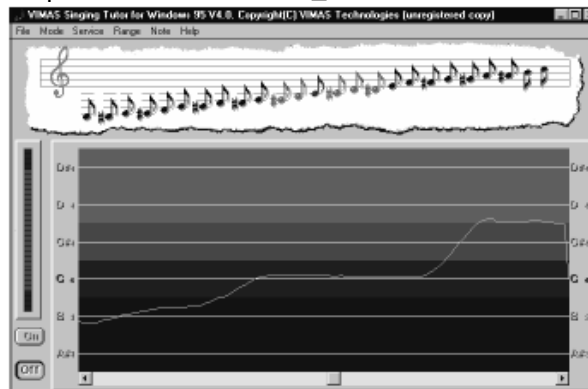
País, año: EU, 1996
Tipo de licencia: Comercial
Campo de aplicación: Análisis de voz hablada y cantada.
Usuarios: Profesores de canto, especialistas del habla.
Funciones: Análisis espectral en tiempo real, electroglotograma, forma de onda, análisis de grabaciones.
Sistema operativo: Windows 95 en adelante. Utiliza un hardware especial (electroglotógrafo), puesto sobre el usuario.
Referencia: <http://www.vocevista.com/index.html>
Pantalla principal:



3. Programas para entrenamiento de la voz cantada

Singing Tutor 4.1.5

País, año: Ucrania, 2007
Tipo de licencia: Comercial. Precio US\$19
Campo de aplicación: Entrenamiento de la voz cantada
Usuarios: Músicos, estudiantes de canto.
Funciones: Estimación de altura de la nota musical cantada.
Sistema operativo: Windows XP
Referencia: http://www.vimas.com/ve_str.htm
Pantalla principal:



Sing & See

País, año:

Australia, 2009

Tipo de licencia:

Comercial. Precios: para estudiantes US\$49, para profesionales US\$99 y el paquete para profesores US\$169 (éste incluye un manual)

Campo de aplicación:

Entrenamiento de la voz cantada

Usuarios:

Músicos, estudiantes de canto.

Funciones:

Estimación de altura de la nota musical cantada. En la versión para profesionales muestra un espectrograma.

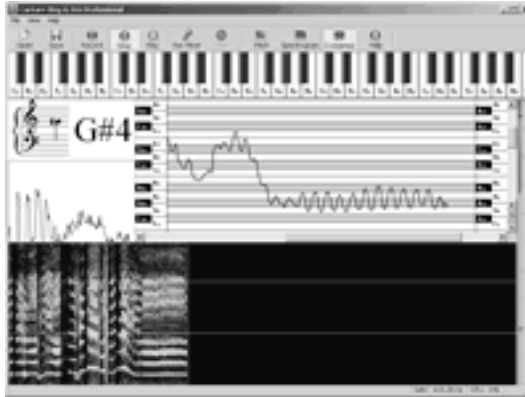
Sistema operativo:

Windows 2000/ XP/ 2003 / Vista

Referencia:

<http://www.singandsee.com/forsingers.php>

Pantalla principal:



Overtone Analyzer

País, año:

Alemania, 2010

Tipo de licencia:

Libre

Campo de aplicación:

Análisis de voz cantada.

Usuarios:

Profesores de canto.

Funciones:

Espectrograma, frecuencias y armónicos, grabaciones.

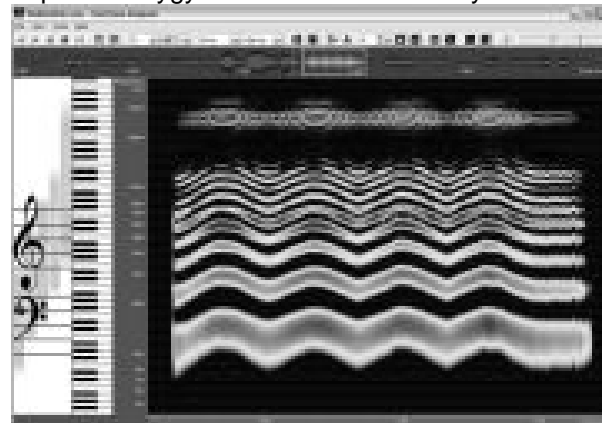
Sistema operativo:

Windows XP, Windows Vista.

Referencia:

<http://www.sygyt.com/en/overtone-analyzer>

Pantalla principal:



APÉNDICE B

CÓDIGO FUENTE DE “EVOCANTO”

Se presenta a continuación un fragmento del código fuente en lenguaje Python del programa *Evocanto*, que contiene principalmente los procedimientos de captura y análisis de la señal de voz. Hacia el final, se omiten unas 15 páginas de código fuente que implementan la interfaz gráfica del programa.

```
#!/usr/bin/python
# -*- coding: cp1252 -*-

from sndobj import *
import wx
import time
import math

class principal(wx.Frame):
    def __init__(self, parent, id, title, size, style):
        wx.Frame.__init__(self, parent, id, title, size =
(575,475), style=wx.DEFAULT_FRAME_STYLE ^ (wx.MAXIMIZE_BOX))
        self.parent = parent
        self.dc = wx.ClientDC(self)
        self.piano = wx.ClientDC(self)
        self.notaPiano = wx.ClientDC(self)
        self.medS1 = wx.ClientDC(self)
        self.medS2 = wx.ClientDC(self)
        self.medCe = wx.ClientDC(self)
        self.medVr = wx.ClientDC(self)
        self.timer = wx.Timer(self, 1)
        self.Bind(wx.EVT_TIMER, self.OnTimer, self.timer)
        self.Bind(wx.EVT_LEFT_DOWN, self.PlayNote)
        self.Bind(wx.EVT_LEFT_UP, self.SilencNote)
        self.tambloq = 32768
        self.highest = 5000
        self.intertime = 743
        self.magniMax = 50
        self.frecmax = 0
        self.indice = 0
        self.tecla = 0
        self.octava = 0
        self.play = False
        self.help = False
        self.data = []
        self.mag = []
        self.frecharm = []
        self.magharm = []
        self.FHz = 0
        self.CreateObjects()
        self.ObjectsWx()

    def CreateObjects(self):
        self.win = HammingTable(self.tambloq,0.5)
        self.inp = SndRTIO(1,SND_INPUT)
        self.sig = SndIn(self.inp, 1)
        self.fft = FFT(self.win,self.sig,1,self.tambloq)
```



```

self.thread = SndThread()
self.threadP = SndThread()
self.thread.AddObj(self.inp, SNDIO_IN)
self.thread.AddObj(self.sig)
self.thread.AddObj(self.fft)

def Analisis(self):
    ind = 0
    self.data = []
    self.mag = []
    self.magV = []
    maxmag = 0
    nmax = int(self.tambloq * self.highest / self.fft.GetSr() + 0.5)
    for n in range(0,self.tambloq/2,1):
        x = self.fft.Output(2*n) * self.tambloq
        y = self.fft.Output(2*n+1) * self.tambloq
        mag = ((x*x) + (y*y))/self.tambloq
        self.magV.append(mag)
        mag = 10 * math.log10(mag + 1)
        self.mag.append(mag)
        if maxmag < mag:
            maxmag = mag
        if n <= nmax:
            ind = 350 * n / nmax
            punto = wx.Point(int(ind) + 1, int(mag*25/20) + 12)
            self.data.append(punto.Get())
    if maxmag < self.magniMax: #no hay nota.
        self.DraWindowSpectro()
        self.DrawPiano()
        self.LimpiaResults()
        self.picture.SetBitmap(wx.Bitmap(INSTALL_PREFIX + 'evomejor.jpg'))
        self.sb.SetStatusText('No hay señal de audio')
    else:
        self.sb.SetStatusText('Ejecutando análisis')
        self.FrecFund()
        self.Harmonics()
        self.Formants()
        self.Vocales()
        self.DrawPiano()
        self.DrawNotaPiano()
        self.DrawSpectro()
        self.LimpiaResults()
        self.PrintResults()

def FrecFund(self):
    self.ajuste = '0'
    nmin = int(100 * self.tambloq / self.fft.GetSr() + 0.5)
    nmax = int(1046 * self.tambloq / self.fft.GetSr() + 0.5)
    nptos = int(10 * self.tambloq / self.fft.GetSr() + 0.5)
    nArm = 6
    maxsuma = 0
    for n in range(nmin,nmax + 1,1):
        suma = 0
        for a in range(1,nArm + 1,1):
            for x in range(-nptos,nptos,1):
                suma = suma + self.mag[a*n+x]

```

```

    if maxsuma < suma:
        maxsuma = suma
        self.frecmax = maxsuma
        self.indice = n

suma = 0
suman = 0
n = self.indice
for a in range(1,nArm + 1,1):
    for x in range(-nptos,nptos + 1,1):
        suma = suma + self.mag[a*n+x]
        suman = suman + self.mag[a*n+x]*(a*n+x)/a
self.FHz = suman/suma * self.fft.GetSr() / self.tambloq
Nc = suman/suma
nptos = int(10 * self.tambloq / self.fft.GetSr() + 0.5)
#Vibrato
suma = 0
suman2 = 0
for a in range(1,2):
    for x in range (-nptos,nptos + 1,1):
        suma = suma + self.magV[a*n+x]
        suman2 = suman2 + self.magV[a*n+x]* math.pow((a*n+x)/a-Nc,2)
self.DeltaFHz = math.sqrt(suman2/suma) * self.fft.GetSr() / self.tambloq
self.FrecVibr = round(self.DeltaFHz,1)
#Nota musical
Rel = self.FHz/440
cents = 1200 * math.log10(Rel + 1e-20) / math.log10(2)
centsC5A = 900
centsC0 = cents + centsC5A + (1200 * 5)
octv = int(math.floor(centsC0/1200))
cents = centsC0 - (1200 * octv)
semitns = int(math.floor(cents/100))
cents = cents - (100 * semitns)
self.difcents = int(math.floor(cents))
if self.difcents>50:
    self.difcents = self.difcents - 100
    semitns = semitns + 1
    if semitns>11:
        semitns = semitns - 12
        octv = octv + 1
self.note = notes[semitns]
self.indnote = str(octv)
self.tecla = semitns
self.octava = octv
if self.difcents > 0:
    self.ajuste = '+' + str(self.difcents)
else:
    self.ajuste = '' + str(self.difcents)

def Harmonics(self):
    self.frecharm = []
    self.magharm = []
    frec = self.FHz
    NH = int(round(self.highest/frec))
    nfund = int(self.tambloq * frec / self.fft.GetSr())
    for i in range(1,NH):

```

```

    freq = self.FHz * i
    n0 = int(self.tambloq * freq / self.fft.GetSr())
    maxmag = self.mag[n0]
    nmax = n0
    for n in range(n0-int(round(nfund/4)),n0+int(round(nfund/4))):
        if maxmag < self.mag[n]:
            maxmag = self.mag[n]
            nmax = n
    self.frecharm.append(nmax * self.fft.GetSr()/self.tambloq)
    self.magharm.append(self.mag[nmax])

def KeyMag(self,i):
    return self.magFormant[i]

def KeyFrec(self,i):
    return self.formant[i]

def Formants(self):
    FindFormant = True
    formant = []
    magFormant = []
    self.formant = []
    self.magFormant = []
    self.sint = []
    indices = []
    count = 0
    min_formant = 200
    if self.magharm[0] > self.magharm[1] and FindFormant == True:
        if self.frecharm[0] > min_formant:
            self.formant.append(self.frecharm[0])
            self.magFormant.append(self.magharm[0])
            self.sint.append(0) #Sintonía de formantes
            indices.append(count)
            count = count + 1
        FindFormant = False
    for i in range(1,len(self.frecharm)-1,1):
        if self.magharm[i] > self.magharm[i+1] and FindFormant == True:
            sumFrecs = (self.frecharm[i-1]*self.magharm[i-1] +
self.frecharm[i]*self.magharm[i] + self.frecharm[i+1]*self.magharm[i+1])
            sumMags = self.magharm[i-1] + self.magharm[i] + self.magharm[i+1]
            promFrecs = sumFrecs/sumMags
            promMags = sumMags/3
            if self.frecharm[i] > min_formant:
                self.formant.append(promFrecs)
                self.magFormant.append(promMags)
                self.sint.append((promFrecs-self.frecharm[i])/self.frecharm[0] * 100) #Sintonía
de formantes
                indices.append(count)
                count = count + 1
            FindFormant = False
        if self.magharm[i] < self.magharm[i+1] and FindFormant == False:
            FindFormant = True
    indices.sort(None,self.KeyMag,True)
    indices = indices[0:5]
    self.FormCant = 0
    for i in range(0,min([len(indices),3])):

```

```

        if self.formant[indices[i]] > 2400:
            self.FormCant = self.formant[indices[i]]
            break
indices.sort(None,self.KeyFrec,False)
formant = self.formant
magFormant = self.magFormant
self.formant = []
self.magFormant = []
for i in range(0,len(indices)):
    self.formant.append(formant[indices[i]])
    self.magFormant.append(magFormant[indices[i]])

def Vocales(self):
    vocal = -1
    F1v = [770,530,320,520,330]
    F2v = [1600,2100,2500,1200,800]
    if len(self.formant)> 1:
        F1 = self.formant[0]
        F2 = self.formant[1]
        # Calcula punto vocálico más próximo a (F1,F2)
        dmin = 1000000
        for v in range(0,5):
            d = math.sqrt(math.pow(F1 - F1v[v],2)+ math.pow(F2 - F2v[v],2))
            if d < dmin:
                dmin = d
                vocal = v
        if dmin > 2000:
            vocal = -1
        if (F1 < 200 or F1 > 1200) or (F2 < 600 or F2 > 3200):
            vocal = -1
    if vocal == -1:
        self.picture.SetBitmap(wx.Bitmap('TractoNormal.jpg'))
    else:
        if vocal == 0:
            self.picture.SetBitmap(wx.Bitmap('TractoA.jpg'))
        if vocal == 1:
            self.picture.SetBitmap(wx.Bitmap('TractoE.jpg'))
        if vocal == 2:
            self.picture.SetBitmap(wx.Bitmap('Tractol.jpg'))
        if vocal == 3:
            self.picture.SetBitmap(wx.Bitmap('TractoO.jpg'))
        if vocal == 4:
            self.picture.SetBitmap(wx.Bitmap('TractoU.jpg'))

```

#Se omiten 15 páginas del código para la interfaz gráfica.

```

if __name__ == "__main__":
    #Inicio del programa
    app = wx.App()
    frame = principal(None,-1,'Evaluación de la Voz en el Canto',size =
(575,475),style=wx.DEFAULT_FRAME_STYLE ^(wx.MAXIMIZE_BOX))
    app.MainLoop()

```

APÉNDICE C

VENTANA DE AYUDA DE “EVOCANTO”

La ventana de ayuda de *Evocanto*, se activa al dar click en el botón de ayuda “?”. Se presenta como un panel (o extensión) a la derecha de la pantalla principal, en el que se muestran los conceptos utilizados en *Evocanto*, como se muestra en la figura C.1, y que son los siguientes:

- Diagramas del tracto vocal
- Espectro de voz cantada
- Formante del cantante
- Vibrato
- Sintonía formante-armónico
- Nota musical y afinación

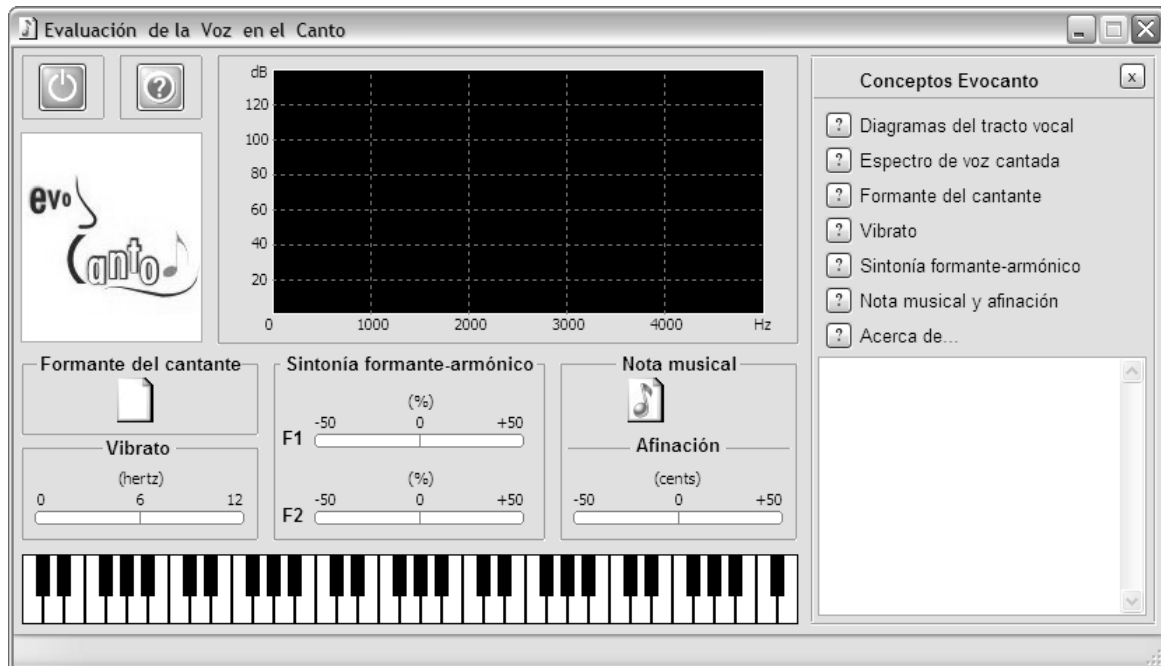


Figura C.1 Panel derecho de Ayuda que contiene los conceptos de Evocanto.

Además de los conceptos utilizados, se incluye un tópico ‘Acerca de...’, donde se indica la versión del programa, los autores y otras características técnicas.

Los conceptos están asociados a botones “?”, que al recibir un click por parte del usuario, despliegan un sub-panel donde se ubica la información correspondiente a cada concepto; en el mismo momento, el título del concepto es resaltado en *negritas*. Por ejemplo, se aprecia en la figura C.2 qué sucede al dar click en “Nota musical y afinación”.

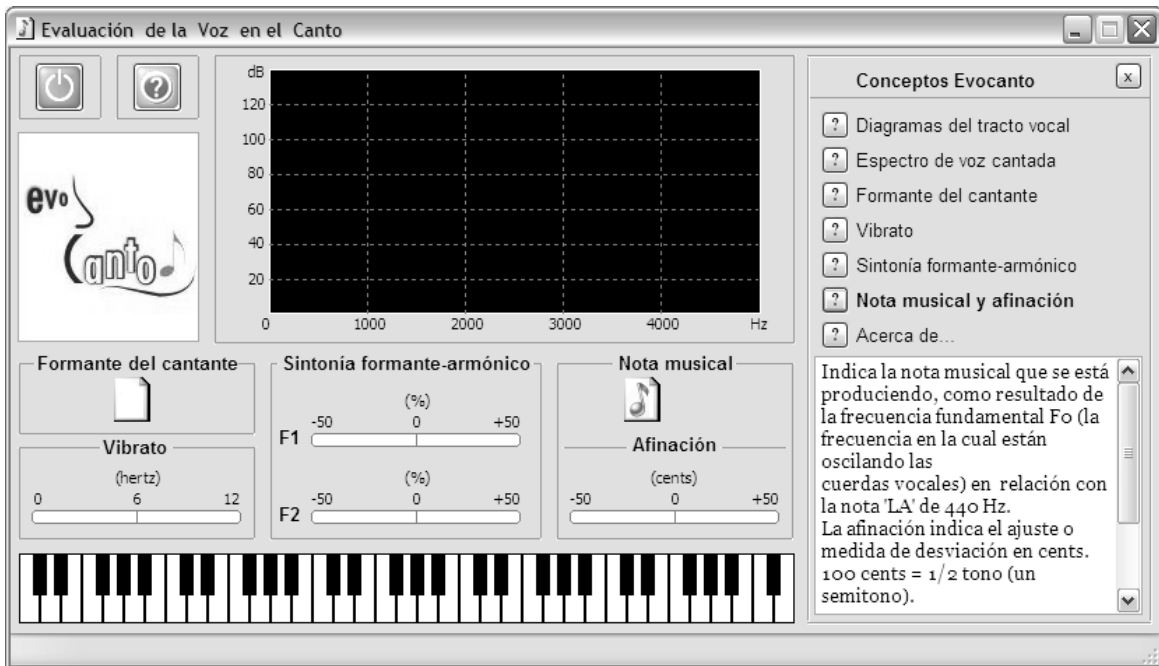


Figura C.2 Despliegue del concepto "Nota musical y afinación" de la ventana de Ayuda de Evocanto.

La descripción de cada concepto se presenta a continuación.

Diagramas del tracto vocal

Para determinar la forma o configuración del tracto vocal, se toma en cuenta la relación entre las formantes F1 y F2. De esa manera se indica la posición horizontal y vertical de la lengua y la abertura de la boca. También se puede determinar de modo aproximado si existe una elongación del tracto vocal al protuir los labios.

Espectro de voz cantada

El espectro es un gráfico que contiene la información de la frecuencia (eje x) en hertz y la intensidad (eje y), en decibeles, de la onda de voz digitalizada.

Se observa en color azul el espectro completo de la voz.

Los armónicos están indicados en color rojo y la envolvente espectral en color amarillo.

Las formantes son picos de resonancia, determinados por la forma y desempeño del tracto vocal.

Formante del cantante.

Cuando las formantes F3, F4 y F5 se acercan entre sí, producen un pico de resonancia entre 2500 y 3000 Hz, llamado "formante del cantante".

Indica que la voz puede destacar por encima del enorme sonido de una orquesta.

Se encuentra en cantantes de registros graves (bajos, barítonos, tenores; a veces contraltos y mezzosopranos) y entrenados profesionalmente.

Se produce al cantar con una confortable posición baja de la laringe.

Sintonía formante-armónico

Cuando se sintonizan las formantes F1 y/o F2 con la frecuencia fundamental, cada uno de los armónicos actúa como una formante. A esto se le denomina "sintonía de formantes con armónicos".

Indica que la voz es resonante.

Se obtiene abriendo mucho la mandíbula.

Se encuentra en todos los cantantes. Pero es un recurso más utilizado por las sopranos.

Si es cero, existe una total sintonía.

Si es negativo, la formante está más abajo que el armónico más cercano.

Si es positivo, la formante está más arriba que el armónico más cercano.

Vibrato

Mide la frecuencia de vibración natural de la voz producida en la glotis (entre las cuerdas vocales), por la acción de un correcto apoyo.

El sistema vocal de un cantante bien entrenado, posee una tendencia natural de producir un vibrato cuya frecuencia se encuentra entre 5 y 7 Hz.

En voces principiantes es menor a 5 Hz.

Más de 7 Hz es un vibrato enfermizo o de voz envejecida.

Es difícil cambiar esta medida de vibrato a voluntad, pues depende del funcionamiento integral del sistema vocal.

Nota musical y afinación

Indica la nota musical que se está produciendo, como resultado de la frecuencia fundamental F0 (la frecuencia en la cual están oscilando las cuerdas vocales) en relación con la nota 'LA' de 440 Hz.

La afinación indica el ajuste o medida de desviación en cents.

100 cents = 1/2 tono (un semitono).

Cuando es cero hay una total afinación; cuando es negativo la nota está baja y cuando es positivo la nota está alta.

Finalmente, para cerrar la ventana de ayuda, basta con dar click en el correspondiente botón "x", o volver a dar click en el botón de ayuda ""?" de la pantalla principal.