



Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

**APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA  
EN MANUAL DE CANTO INTERACTIVO, BASADO EN LA  
TÉCNICA ALEMANA**

**Marco Vinicio Gómez Salazar**

Asesorado por el Ing. Miguel Antonio Morales Guillen

Guatemala, junio de 2009



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA  
EN MANUAL DE CANTO INTERACTIVO, BASADO EN LA  
TÉCNICA ALEMÁNA**

TRABAJO DE GRADUACIÓN

PRESENTADO A JUNTA DIRECTIVA  
DE LA FACULTAD DE INGENIERÍA  
POR

**MARCO VINICIO GÓMEZ SALAZAR**

ASESORADO POR EL ING. MIGUEL ANTONIO MORALES GUILLEN

AL CONFERÍRSELE EL TÍTULO DE  
**INGENIERO EN CIENCIAS Y SISTEMAS**

GUATEMALA, JUNIO DE 2009



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA  
FACULTAD DE INGENIERÍA



**NÓMINA DE JUNTA DIRECTIVA**

Decano	Ing. Murphy Olympo Paiz Recinos
Vocal I	Inga. Glenda Patricia García Soria
Vocal II	Inga. Alba Maritza Guerrero de López
Vocal III	Ing. Miguel Ángel Dávila Calderón
Vocal IV	Br.. José Milton De León Bran
Vocal V	Br. Isaac Sultán Mejía
Secretaria	Inga. Marcia Ivónne Véliz Vargas

**TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO**

Decano	Ing. Murphy Olympo Paiz Recinos
Examinadora	Inga. Virginia Victoria Tala Ayerdi
Examinador	Ing. Pedro Pablo Hernández Ramírez
Examinador	Ing. Cesar Augusto Fernández Cáceres
Secretaria	Inga. Marcia Ivónne Véliz Vargas



## HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

### **APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA EN MANUAL DE CANTO INTERACTIVO, BASADO EN LA TÉCNICA ALEMÁNA,**

tema que me fuera asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, en agosto de 2008.



Marco Vinicio Gómez Salazar



Guatemala, 11 de Mayo de 2009

Ingeniero  
Carlos Azurdia  
Escuela de Ingeniería en Ciencias y Sistemas  
Facultad de Ingeniería  
Universidad de San Carlos de Guatemala

Ingeniero Azurdia.

Al saludarle me dirijo a usted para informarle que he tenido a bien asesorar el trabajo de graduación titulado: **“APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA EN MANUAL DE CANTO INTERACTIVO BASADO EN LA TECNICA ALEMANA”**, desarrollado por el estudiante **Marco Vinicio Gómez Salazar** carné **200112583**, y encontrándolo satisfactorio en su contenido y resultados, me permito dar aprobación al mismo sugiriendo para el tramite, en el entendido que el Autor y Asesor somos responsables del contenido y conclusiones del trabajo.

Sin otro particular me suscribo.

Atentamente,



*Miguel Antonio Morales Guillen*  
Ingeniero en Ciencias y Sistemas  
Colegiado No. 8460

Ing. Miguel Antonio Morales Guillen  
Colegiado No. 8460.

**ASESOR**





Universidad San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 25 de Mayo de 2009

Ingeniero  
**Marlon Antonio Pérez Turk**  
Director de la Escuela de Ingeniería  
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **MARCO VINICIO GÓMEZ SALAZAR**, titulado: **"APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA EN MANUAL DE CANTO INTERACTIVO BASADO EN LA TÉCNICA ALEMANA"**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,

  
**Ing. Carlos Alfredo Azurdia**  
Coordinador de Privados  
y Revisión de Trabajos de Graduación





E  
S  
C  
U  
L  
A  
  
D  
E  
  
C  
I  
E  
N  
C  
I  
A  
S  
  
S  
I  
S  
T  
E  
M  
A  
S

UNIVERSIDAD DE SAN CARLOS  
DE GUATEMALA



FACULTAD DE INGENIERÍA  
ESCUELA DE CIENCIAS Y SISTEMAS  
TEL: 24767644

*El Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, de trabajo de graduación titulado "APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA EN MANUAL DE CANTO INTERACTIVO, BASADO EN LA TÉCNICA ALEMANA", presentado por el estudiante MARCO VINICIO GÓMEZ SALAZAR, aprueba el presente trabajo y solicita la autorización del mismo.*

**"ID Y ENSEÑAD A TODOS"**

*Ing. Marlon Antonio Pérez Turk*

*Director, Escuela de Ingeniería Ciencias y Sistemas*



*Guatemala, 10 de junio 2009*



Universidad de San Carlos  
de Guatemala



Facultad de Ingeniería  
Decanato

Ref. DTG.205.09

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al trabajo de graduación titulado: **APLICACIÓN DE ANÁLISIS ESPECTRAL DE LA VOZ HUMANA EN MANUAL DE CANTO INTERACTIVO, BASADO EN LA TÉCNICA ALEMANA**, presentado por el estudiante universitario **Marco Vinicio Gómez Salazar**, procede a la autorización para la impresión del mismo.

IMPRÍMASE.

A handwritten signature in black ink, consisting of a large loop at the top and several vertical strokes below, positioned over the printed name and title.

Ing. Murphy G. Paiz Recinos  
DECANO



Guatemala, junio de 2009

/cc



## **AGRADECIMIENTOS A:**

<b>Dios</b>	Por hacerme sentir amado y darme al Espíritu Santo como fuente de fortaleza, sabiduría y templanza en toda mi vida.
<b>María Auxiliadora, San Juan Bosco y Domingo Savio</b>	Por ser mis amigos de siempre y enseñarme a seguir a Dios, por medio de su testimonio de vida
<b>Mis padres: Sandra Elizabeth Salazar de Gómez y Lic. Marco Vinicio Gómez Campos</b>	Por darme su amor, ejemplo, apoyo y la mejor herencia del mundo la fe en Dios.
<b>Mi hermano: Edgar Josué Gómez Salazar</b>	Por todo su amor, y compañía en este viaje de la vida y todas sus estaciones.
<b>El Colegio Salesiano Don Bosco.</b>	Por formarme como un buen cristiano y un honrado ciudadano.
<b>La Universidad de San Carlos de Guatemala</b>	Por su historia y presente, por las buenas y malas que han hecho de mí lo que soy como profesional
<b>Mi tutor y amigo</b>	Ing. Miguel Antonio Morales Guillen



**Mi tío Diego Web**

Por su apoyo en uno de los momentos  
mas difíciles de mi carrera.

**Mi tío Carlos Hernández**

Por su amistad, apoyo y cariño.  
(D.E.P.)



## **ACTO QUE DEDICO A:**

<b>Dios</b>	Padre, Hijo y Espíritu Santo.
<b>Mis padres</b>	Sandra Elizabeth Salazar de Gómez y Lic. Marco Vinicio Gómez Campos
<b>Mi hermano:</b>	Edgar Josué Gómez Salazar
<b>A mis abuelitas(os)</b>	Marta de Salazar Efraín Salazar (D.E.P) Yolanda de Gómez Víctor Manuel Gómez (D.E.P) Juan José Corado (D.E.P)
<b>Mis tíos (as)</b>	Sandra, Paty, Claudia, Aury, Victor, Diego, Carlos (D.E.P), Rony (D.E.P) y Oscar.
<b>Mis primos(as)</b>	Vivian, Susy, Rita, Michelle, Paula, Marcela, Maria de los Ángeles (D.E.P), Javier, Alejandro, Aleida, Heidy, Ingrid y Cristina.
<b>Mi novia</b>	Arely Carolina Cano López y su familia gracias por su cariño y amistad.
<b>Mis amigos</b>	De toda la vida, los nuevos amigos y los que están por venir, porque “sin duda alguna, la vida es más fácil con ustedes”.



**Mis padrinos**

Sandra, Edwin, Luis y Olga (D.E.P)

**Mis ahijados(as)**

Samanta, Josué y Erick.

**La obra de  
Don Bosco.**

Hermanos de la Promoción LI, Claustro de Maestros y alumnos del CSDB, Grupo Génesis, Parroquia Sagrado Corazón de Jesús.

**Mi país**

Guatemala y toda su gente que lucha día a día por ser mejor y dar un futuro diferente a sus hijos.

**La vida**

Que me ha dado tanto y parece un suspirar cuando se vive a plenitud.



## ÍNDICE GENERAL

<b>ÍNDICE DE ILUSTRACIONES</b>	<b>V</b>
<b>GLOSARIO</b>	<b>VII</b>
<b>RESUMEN</b>	<b>XI</b>
<b>OBJETIVOS</b>	<b>XIII</b>
<b>INTRODUCCIÓN</b>	<b>XV</b>
<b>1. EL ARTE DE CANTAR Y LA TÉCNICA DE CANTO ALEMANA</b>	<b>1</b>
1.1 El arte de cantar	1
1.2 Técnica de canto alemana	3
1.2.1 Precalentamiento	4
1.2.1.1 Relajación del cuello	5
1.2.1.2 Calentando las cuerdas vocales	6
1.2.1.3 Arpeggios bajos	6
1.2.1.4 Relajación de los músculos faciales	7
1.2.2 Vocalización	8
1.2.2.1 Técnica de respiración diafragmática	9
1.2.2.2 Ejercicio de respiración completa	10
1.2.2.3 Ejercicio de respiración abdominal	11
1.2.3 Desarrollo de la afinación	12
1.2.3.1 Método de afinación 1	12
1.2.3.2 Método de afinación 2	12
<b>2. PROCESO DE PRODUCCIÓN Y CAPTURA DE LA VOZ HUMANA</b>	<b>15</b>
2.1 El aparato fonador	15

2.2 El aparato auditivo	20
2.3 El sonido	23
2.4 Características musicales del sonido	25
2.4.1 Frecuencia	25
2.4.2 Amplitud	26
2.4.3 Forma de onda	26
2.5 Archivos de audio	27
<b>3. ANÁLISIS ESPECTRAL DE LA VOZ HUMANA</b>	<b>31</b>
3.1 Análisis espectral	31
3.1.1 Transformada de Fourier	33
3.1.2 DTF y FFT	34
3.2 Propiedades de la voz digitalizada	37
3.2.1 Ancho de banda y frecuencia de muestreo	38
3.2.2 Largo del bloque de análisis	39
3.2.3 Enventanado	40
3.2.4 Resolución espectral	40
<b>4. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE MANUAL DE CANTO INTERACTIVO, BASADO EN LA TÉCNICA DE CANTO ALEMANA</b>	<b>43</b>
4.1 Análisis y diseño	43
4.1.1 Requerimientos funcionales	43
4.1.2 Requerimientos no funcionales	45
4.1.3 Herramientas utilizadas	47
4.1.3.1 Cool Edit Pro 2	47
4.1.3.2 Matlab 7.6.0. R2008a	47
4.2 Arquitectura	48
4.2.1 Micrófono	48
4.2.2 Tarjeta de audio	48

4.2.3 Interfaz gráfica de usuario	49
4.2.4 Funciones de I/O de Matlab	49
4.2.5 Archivos de audio	49
4.2.6 Análisis espectral con la función FFT de Matlab	49
4.3 Paradigma de programación	50
4.3.1 Programación orientada a eventos	51
4.4 Implementación de manual de canto interactivo, basado en la técnica de canto alemana	51
4.4.1 Grabación de lecciones	51
4.4.2 Grabación de ejercicios	54
4.4.3 Aplicación de transformada de Fourier en archivos de audio	56
4.4.4 Comparación de vectores de frecuencias fundamentales	59
4.4.5 Mostrar e interpretar resultados	62
<b>CONCLUSIONES</b>	<b>65</b>
<b>RECOMENDACIONES</b>	<b>67</b>
<b>BIBLIOGRAFÍA</b>	<b>69</b>
<b>ANEXOS</b>	<b>71</b>



# ÍNDICE DE ILUSTRACIONES

## FIGURAS

1	Ejercicio de vocalización “5 vocales”	9
2	Postura para ejercicio de respiración completa	11
3	Aparato fonador	15
4	Funcionamiento del aparato fonador	20
5	Aparato auditivo	21
6	Gráfica de la función $x(t) = A\cos(\omega t + \varphi)$	24
7	Áreas de audición	29
8	Gráfica de espectro de señal sonora	32
9	Gráfica de la señal $\tilde{x}[n]$	35
10	DFT de la señal muestreada $\tilde{x}[n]$	36
11	Espectrograma de señal de voz	39
12	Arquitectura del sistema	50
13	Gráfica de señal de audio de tiempo versus amplitud	52
14	Grabación de ejercicio	55
15	Gráfica de resultados de análisis de ejercicio maestro y alumno	63
16	Ubicación del archivo instalador manual_canto.exe	72
17	Archivos utilizados para la ejecución de la aplicación	73

18	Pantalla inicial de instalación de MCRInstaller.exe	74
19	Aplicación “Manual de Canto Interactivo”	74
20	Entorno de trabajo “Manual de Canto Interactivo”	75
21	Frame maestro y alumno	76
22	Ubicación botones de control	77
23	Gráfica de la lección 1	79
24	Gráfica de grabación de ejercicio realizado por el alumno	80

## **TABLAS**

I	Parámetros típicos de la digitalización de sonido	27
II	Niveles de dificultad en el análisis de ejercicios	61

## GLOSARIO

<b>Arpeggio</b>	El arpeggio es una manera de ejecutar los tonos de un acorde: en vez de tocarlos de manera simultánea, se hacen oír en sucesión rápida, generalmente del más grave al más agudo.
<b>Ataque cómodo</b>	Ejecución de la voz cantada en frecuencias cómodas para el cantante.
<b>Bits</b>	Bit es el acrónimo de <i>Binary digit</i> . (dígito binario). Un bit es un dígito del sistema de numeración binario. La Real Academia Española (RAE) ha aceptado la palabra bit con el plural bits.
<b>Canal</b>	Es el medio de transmisión por el que viajan las señales portadoras de la información que pretenden intercambiar emisor y receptor.
<b>Color</b>	Básicamente, es la técnica empleada o bien es la conducta vocal, la que determina el color del canto, siendo éste claro u oscuro. Dentro del color, tenemos la eufonía, siendo ésta, el matiz que el cantante emplea en la emisión vocálica, así pues, un cantante puede presentar una eufonía clara u oscura.
<b>Compas</b>	Es la entidad métrica musical, compuesta por varias unidades de tiempo (como la negra o la corchea).
<b>Dominio</b>	Conjunto de valores para los que una determinada función matemática está definida,
<b>Fortissimo</b>	Fortissimo (del italiano "muy fuerte") es un término utilizado en notación musical para indicar una determinada intensidad en el sonido, es decir, un determinado matiz.

<b>Frame</b>	En lo referente a la informática, se le llama <i>frame</i> en inglés a lo que en castellano se conoce como marco o cuadro.
<b>Hamming</b>	Las ventanas son funciones matemáticas usadas con frecuencia en el análisis y el procesamiento de señales, para evitar las discontinuidades al principio y al final de los bloques analizados
<b>Hertz(HZ)</b>	El hertz o hercio es la unidad de frecuencia del Sistema Internacional de Unidades. Un hertz representa un ciclo por cada segundo, entendiendo <i>ciclo</i> como la repetición de un evento.
<b>If</b>	En programación, una sentencia condicional es una instrucción o grupo de instrucciones que se pueden ejecutar o no en función del valor de una condición.
<b>Metrónomo</b>	El metrónomo es un aparato utilizado en música para medir el tempo (la velocidad de la música).
<b>Midi</b>	Son las siglas de <i>Musical Instrument Digital Interface</i> (Interfaz Digital de Instrumentos Musicales). Se trata de un protocolo industrial estándar que permite a las computadoras, sintetizadores, secuenciadores, controladores y otros dispositivos musicales electrónicos comunicarse y compartir información para la generación de sonidos.
<b>Minnesänger</b>	Trovadores alemanes.
<b>Octava Baja:</b>	Escala de las siete notas musicales en frecuencia baja, iniciando en do y terminando nuevamente en do, pero en una frecuencia más aguada al do inicial.
<b>Pianissimo</b>	Pianissimo (del italiano <i>muy suave</i> ) es un término utilizado en música para indicar una determinada intensidad en el sonido, es decir, un determinado matiz.
<b>Plotear</b>	Es la acción de trazar un gráfico, gráficos vectoriales o dibujos lineales: planos, dibujos de piezas, etc.

<b>Rangos altos</b>	Vibraciones más rápidas de las cuerdas vocales.
<b>Rangos bajos</b>	Vibraciones más bajas de las cuerdas vocales.
<b>Troubadours</b>	Trovadores Italianos.
<b>Trouvéres</b>	Trovadores franceces.
<b>Vector</b>	En programación, un array es un conjunto o agrupación de variables del mismo tipo cuyo acceso se realiza por índices.
<b>Vibrato</b>	El vibrato (it. vibrato = vibración) es una variación periódica de la frecuencia y / o amplitud de un sonido. En el canto se produce espontáneamente a través de un tremor nervioso en diafragma y / o laringe.
<b>Vocalizar</b>	Vocalizar es jugar con una melodía sencilla pero que abarque por lo menos media escala e ir subiéndola de tono.
<b>Wave</b>	WAV (o WAVE), apócope de <i>WAVEform audio format</i> , es un formato de audio digital normalmente sin compresión de datos desarrollado y propiedad de Microsoft y de IBM que se utiliza para almacenar sonidos en el PC, admite archivos mono y estéreo a diversas resoluciones y velocidades de muestreo, su extensión es .wav.
<b>While</b>	El Bucle while es una estructura de la mayoría de los lenguajes de programación estructurados cuyo propósito es repetir un bloque de código mientras una condición se mantenga verdadera.



## RESUMEN

Actualmente, el aprender un arte se ha convertido en un asunto de inversión de tiempo y gran cantidad de dinero. En la búsqueda de una solución a este problema, se orienta este proyecto de fin de carrera al desarrollo de un Manual Interactivo introductorio al canto, el cual integrará módulos de teoría, ejemplos y ejercicios extraídos directamente del Método Alemán de Canto.

Esta aplicación recibirá la voz humana como una señal acústica y la transformará en señales análogas, por medio de un micrófono. Luego estas señales serán convertidas en un conjunto de señales digitales para ser medidas y comparadas.

El valor agregado de esta aplicación radica en los módulos de ejercicios, en los cuales el alumno podrá escuchar una grabación con un ejemplo de vocalización, que este deberá repetir para luego ser evaluado. Para la evaluación de los ejercicios se utilizará el análisis espectral de las señales sonoras, tanto del audio de muestra como del audio grabado por el aprendiz. Además el alumno podrá escuchar su voz en el ejercicio grabado y observar su puntuación para saber en qué aspectos mejorar y medir su nivel de avance.



## OBJETIVOS

### **General:**

Brindar a la sociedad, un método que permita realizar el proceso enseñanza-aprendizaje del canto, de manera interactiva, económica y sencilla. Así mismo, desarrollar una aplicación que dará al investigador la oportunidad de ampliar sus conocimientos en el proceso de grabación y análisis de la voz humana, por medio de técnicas de transformación analógicas y digitales.

### **Específicos:**

1. Proporcionar al usuario una retroalimentación de su nivel de avance en el aprendizaje del canto, por medio de ejercicios de evaluación de la entonación y el ritmo con 85 % de exactitud.
2. Desarrollar tres ejercicios de evaluación de la entonación, vocalización y ritmo.
3. Realizar comparaciones basadas en el análisis espectral de la voz humana utilizando la Transformada de Fourier, en grabaciones con una duración máxima de 30 segundos.
4. Conocer dos de los conceptos básicos del sonido digital: frecuencia y amplitud.



## INTRODUCCIÓN

Actualmente, existen pocas entidades públicas que se dedican a la enseñanza del canto, tal es el caso del Conservatorio Nacional de Música, en el cual los costos son mínimos, pero se requiere de gran cantidad de tiempo para poder asistir a los curso diariamente durante cinco años y la atención no es personalizada.

En contraparte, encontramos las academias privadas, en las cuales la atención es bastante buena y personalizada, pero el estudio dentro de éstas es bastante oneroso y esto provoca que la oportunidad de practicar este arte únicamente sea para cierto grupo de alumnos de nivel económico medio alto y alto.

Llevar el arte del canto a alumnos de todos los estratos socioeconómicos de Guatemala, es la motivación principal de este proyecto, aunado a un deseo de investigación en el cual se demuestre una nueva o muy poco desarrollada aplicación del análisis espectral de la voz humana, utilizando la técnica de Transformada de Fourier que no es más que la representación en sumas de senos y cosenos de una señal periódica dada.

Cabe mencionar que los resultados del presente proyecto de fin de carrera, se vieron afectados por la tecnología y algoritmos utilizados, ya que se debía analizar gran cantidad de datos. Agregado a esto se debe tomar en cuenta el impacto de las diferentes variables que afectan un proceso de grabación, reproducción y análisis de audio.

A pesar de lo antes expuesto, los resultados luego de varios meses de investigación fueron positivos, ya que se logró realizar la síntesis y comparación de frecuencias fundamentales en un tiempo determinado en diferentes señales de audio emitidas por el ser humano durante el proceso de emisión de la voz cantada.

# 1 EL ARTE DE CANTAR Y LA TECNICA DE CANTO ALEMANA

## 1.1 El arte de cantar

El arte del canto nació con el hombre mismo, con su primera expresión vocal. En sus orígenes fue una forma más elevada del lenguaje, probablemente inspirada por el culto primitivo. Hay incluso quienes afirman que el canto existió antes del lenguaje hablado, así como existe en especies inferiores al hombre, por ejemplo, las aves.

Más adelante el canto respondió a las necesidades de las religiones y las estéticas, condicionadas naturalmente por ejemplo, por diferentes lenguas, las que llevaron a distintas maneras de emitir la voz (nasalización y elevación artificial de la laringe en las culturas del Oriente Medio).

En la antigüedad mediterránea, el arte del canto tuvo influencia en la retórica; en Grecia los discursos debían ser rendidos en un determinado tono. Para la tragedia y comedia griegas se necesitaban cantantes formados, que junto con el drama ofrecían secciones cantadas.<sup>1</sup>

---

1.Hanns Stein. El arte de Cantar: Su Dimensión Cultural y Pedagógica. Disponible en: <http://www.cantolirico.com/content/view/368/>

El aporte más importante de la Iglesia Católica al arte del canto es, posiblemente, el canto litúrgico y su desarrollo polifónico en siglos posteriores, antes de concluir la Edad Media. A fines de este mismo período, en Europa, aparece un tipo de canto profano que se podría denominar de arte y que practican los *troubadours*, *trouvères* y *minnesänger*. La improvisación libre del siglo XVI sentó las bases técnico-vocales para el amplio y diferenciado desarrollo que tuvo el arte del canto a partir de 1600, con el florecimiento de la ópera, el oratorio, la cantata y el aria. En ese período, en el cual el canto tuvo un desarrollo musical tan grande, la principal preocupación era la expresión y la comprensibilidad del texto.

El virtuosismo vocal se desarrolló a la par con el instrumental, y sus grandes cultores fueron los castrati con su enorme habilidad en la ejecución de coloraturas. Recién el nuevo dramatismo del clasicismo restauró la relación original entre música y declamación y reemplazó al castrato por el cantante dramático, así como permitió la integración de las mujeres, que fueron excluidas del canto desde el siglo VII, por la Iglesia Católica.

En el siglo XIX apareció la canción artística (Lied, Kunstlied), la que con sus exigencias técnicas y expresivas enriqueció el arte del canto. En el arte vocal francés, en todas las épocas, el acento estaba puesto en la declamación de la palabra.<sup>2</sup>

---

<sup>2</sup> Hanns Stein. El arte de Cantar: Su Dimensión Cultural y Pedagógica. Disponible en: <http://www.cantolirico.com/content/view/368/>

La Escuela Belcantística italiana, en cambio, enseñaba el libre desarrollo del melo vocal, en detrimento de lo declamatorio. Los alemanes y otras naciones centroeuropeas encontraron un cierto equilibrio entre los dos extremos. Posteriormente, incluso maestros como Verdi y Puccini exigían unidad orgánica entre palabra y sonido musical.<sup>3</sup>

Hoy en día aunque se siguen poniendo en práctica las escuelas de canto francesa e italiana, la técnica alemana ha creado un equilibrio de ambas escuelas es por ello que en el presente trabajo de fin de carrera utilizaremos como base de estudio la técnica de la Escuela de Canto Alemana

## **1.2 Técnica de canto alemana**

Indistintamente de la técnica utilizada al cantar es vital que al utilizar el aparato fonador se utilice un sistema o técnica adecuada para no poner en riesgo la salud de los órganos que intervienen en la ejecución de la voz cantada. La técnica de canto alemana brinda diversidad de ejercicios que mejoran rápida y notoriamente la calidad de la ejecución de la voz al cantar, abarcando todos los aspectos necesarios para una buena interpretación como lo son: respiración, afinación, vibrato, volumen, y color y ritmo.

Esta técnica permite ser aplicada a los diferentes estilos y ritmos de la música contemporánea y clásica sin que el alumno se deba orientar directamente al estilo de canto que dio origen a la técnica.

---

<sup>3</sup> Hanns Stein. El arte de Cantar: Su Dimensión Cultural y Pedagógica. Disponible en: <http://www.cantolirico.com/content/view/368/>

La técnica de canto alemana ha sido difundida en el internet gracias al “Método Alemán de Canto” creado por Daniel García y es de acá de donde se extraen algunos términos que nos ocuparan en el presente informe.

### **1.2.1 Pre calentamiento**

Como el cuerpo de un atleta, la voz ha de ser pre calentada suavemente para que trabaje con el mejor rendimiento y no se vea dañada. Si canta con energía a partir de un estado “frío”, especialmente en la parte superior de su voz, producirá con toda seguridad un sonido áspero. Diez minutos de pre calentamiento lento, tranquilo y suave serán suficientes. Tararé un poco, utilice primero la octava baja, practique el ataque cómodo, la apertura de garganta, el control respiratorio y las notas sostenidas.<sup>4</sup>

Los cantantes deben comenzar calentando el cuerpo entero con varios ejercicios físicos. Esto ayuda a aliviar la tensión muscular que interfiere con la producción vocal, así como estimular la respiración profunda, la cual es necesaria para un buen soporte de la voz.

Los músculos de articulación que incluyen el cuello, la mandíbula, lengua, labios y el paladar suave pueden relajarse con ejercicios apropiados, los cuales pueden ayudar para activar el fluido de aire del cantante.<sup>5</sup>

---

<sup>4</sup> Marta Bellomo. Clínicas de Canto: La importancia de la práctica. Disponible en: [http://www.escueladejazz.com.ar/clinicas\\_canto\\_06.htm](http://www.escueladejazz.com.ar/clinicas_canto_06.htm)

<sup>5</sup> *Tips* Lección (1-4) Disponible en: <http://dannyrdz7.tripod.com/id8.html>

### **1.2.1.1 Relajación del cuello**

Estando de pie bien erguido y con los brazos a cada lado del cuerpo, se debe hacer rotar la cabeza sobre los hombros hacia la derecha y hacia la izquierda muy lentamente y con los ojos cerrados.

Es muy importante que la respiración sea rítmica y lenta para que la relajación del cuello sea completa. Se debe exhalar el aire mientras la cabeza desciende e inhalar cuando asciende hasta su posición normal.

Ahora con la vista al frente debe ladear la cabeza hasta que la oreja derecha se apoye sobre el hombro derecho, siempre viendo hacia delante, sentirá como se estira el lado izquierdo del cuello, mantenga esa posición aproximadamente 30 segundos y luego repita lo mismo hacia el lado opuesto.

Una vez terminada la rutina anterior se coloca de nuevo la vista al frente y se rota la cabeza hacia atrás lo más que se pueda dejándola en esa posición unos 15 segundos.

Asegúrese de no perder la postura de la espalda, debe estar lo más erguida posible, luego se vuelve a la posición inicial y se inclina cabeza hacia delante tratando de tocar el pecho con el mentón. Mientras hace esto, debe llevar los hombros ligeramente hacia atrás, no debe ser brusco al hacer este ejercicio debe realizarlo suavemente pero con firmeza.<sup>6</sup>

---

<sup>6</sup> Marcelo Roascio. Lecciones de Canto..Disponible en:  
<http://www.guitarraonline.com.ar/canto2.htm>

### **1.2.1.2 Calentando las cuerdas vocales.**

Es sabio comenzar vocalizando en el rango más confortable de la voz, el rango medio, y gradualmente ir expandiéndose hacia rangos más altos y rangos más bajos, mientras el aparato fonador no sufra un sobreesfuerzo.

Se debe vocalizar diciendo “mmm” continuo y en una tonalidad media y cómoda, descansando la lengua en el maxilar inferior, con la punta detrás de los dientes.

Luego se debe decir “me” – “me” – “me” – “me” atribuyendo una nota más alta a cada silaba.<sup>7</sup>

### **1.2.1.3 Arpeggios bajos**

A este ejercicio se le denomina arpeggios bajos ya que se realizan con notas alternadas pertenecientes a una escala determina.

Continuando con el ejercicio “mmm” ahora este ejercicio se debe realizar de forma ascendente, llegando al tono más alto que se pueda obtener sin forzar la voz. Este ejercicio se debe repetir cinco veces, ya que el objetivo primordial de este es hacer subir y bajar “la nuez”, esto produce un trabajo en los músculos que intervienen en la generación de la voz.<sup>8</sup>

---

<sup>7</sup> *Tips* Lección (1-4) Disponible en: <http://dannyrdz7.tripod.com/id8.html>

<sup>8</sup> . Marcelo Roascio. Lecciones de Canto..Disponible en:  
<http://www.guitarraonline.com.ar/canto2.htm>

#### **1.2.1.4 Relajación de los músculos faciales**

No solo hay que trabajar los músculos de la garganta sino también los del rostro, pues estos tienen un papel preponderante en canto, ya que son los encargados de producir no solo las expresiones faciales sino también ayudar a que la voz tome un color y un timbre especial.

El estiramiento de los músculos de la cara debe posibilitar una relajación de las zonas musculares tensas. Como consecuencia de una tensión general de una forma de hablar económica, existen a menudo en la zona de los músculos faciales y masticadores tensiones crónicas, que afectan el timbre de voz.

Se debe realizar este ejercicio sentado, con las espalda recta. Colocando las manos sobre la frente y con las yemas de los dedos haciendo algo de presión, estire lentamente la frente desde el centro hacia las sienes, este masaje se debe repetir 10 veces.

- Luego con las yemas de los dedos índices y mayores apoyados sobre la zona superior de la nariz, estire los músculos en dirección a las mejillas, este masaje se debe repetir 10 veces.
- El siguiente paso es colocar las yemas de los dedos en ambos pómulos y estirar los músculos hacia la barbilla, este masaje se debe repetir 10 veces.<sup>9</sup>

---

<sup>9</sup> Dra. Iris Eicher. Programa. Básico para terapia de voz. Disponible en: <http://www.intervoice.de/sp/basis2.html>

- Luego con los dedos índices puestos entre los labios superior y la nariz estire los músculos presionando con las yemas de los dedos tirando hacia las comisuras de los labios, este masaje se debe repetir 10 veces.
- Por último, se deben colocar las yemas de los dedos índices entre los labios y la barbilla y estirar esos músculos tirando en forma horizontal hacia las comisuras de los labios, este masaje se debe repetir 10 veces.<sup>10</sup>

## 1.2 Vocalización

Vocalizar es jugar con una melodía sencilla pero que abarque por lo menos media escala e ir subiéndola de tono. La vocalización nos sirve para dos cosas, mejorar la entonación y mejorar la dicción.

Podemos decir que la palabra vocalización se deriva de vocal. Las vocales son los sonidos básicos de cualquier idioma y son los sonidos en los que haremos los cambios de tono ver figura 1.

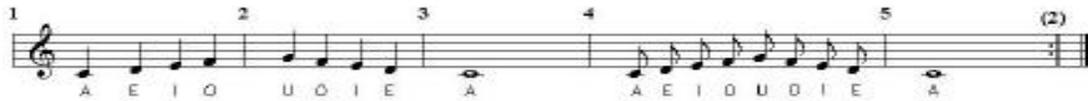
Cada una de estas vocales debe producirse con una posición específica de nuestra boca, ya que por ejemplo, para la vocal “a” nuestra boca deberá estar abierta, mientras que para la “u” debemos tenerla casi cerrada; aunque como ya vimos, debemos conservar el interior de la boca muy grande.<sup>11</sup>

---

<sup>10</sup> Dra. Iris Eicher. Programa. Básico para terapia de voz. Disponible en:  
<http://www.intervoice.de/sp/basis2.html>

<sup>11</sup> Moisés León. Vocalización y Respiración. Disponible en:  
<http://www.gruposmusicalesparroquiales.org/tecnica/Respiracionyvocalizacion.pdf>.

**Figura 1. Ejercicio de vocalización cinco vocales**



Fuente: De León Moisés. Manual de Respiración y Vocalización Pastoral Grupos Parroquiales Juveniles Arquidiócesis Primada de México.

Pag.2

### **1.2.2.1 Técnica de respiración diafragmática.**

Saber administrar en flujo de aire durante una sesión de canto es una de las técnicas más importantes.

Las características dinámicas y sonoras de cada nota que producimos están directamente relacionadas con esta técnica que, si utilizamos mal, puede traer deficiencias realmente notorias a la hora de cantar.

Es importante recordar que el aparato vocal es un instrumento de viento y como tal el flujo de aire lo controla todo. Un flujo parejo y constante de aire producirá una nota continua y controlada además de agradable.

Esta técnica de canto no tiene a la respiración como elemento principal, pero es necesario practicarla porque este método va dirigido a personas con diferentes capacidades respiratorias y será conveniente que utilicen algunas técnicas desarrolladas para mejorarlas.<sup>12</sup>

---

<sup>12</sup>Tips Lección (1-4) Disponible en: <http://dannyrdz7.tripod.com/id8.html>

La respiración en el canto es tan importante como en el habla. Cuando el ser humano habla, inhala y exhala de manera inconsciente, pues entonces debe hacerse lo mismo cuando se canta, debe ser natural y espontáneo. Puede ser que aun cuando se habla se utilice más aire del necesario, lo cual lleva a que se genere tensión en el cuello y se fuerce las cuerdas vocales pudiéndose producir una disfonía, para ello hay diferentes ejercicios que pueden ayudar a controlar mejor la respiración.<sup>13</sup>

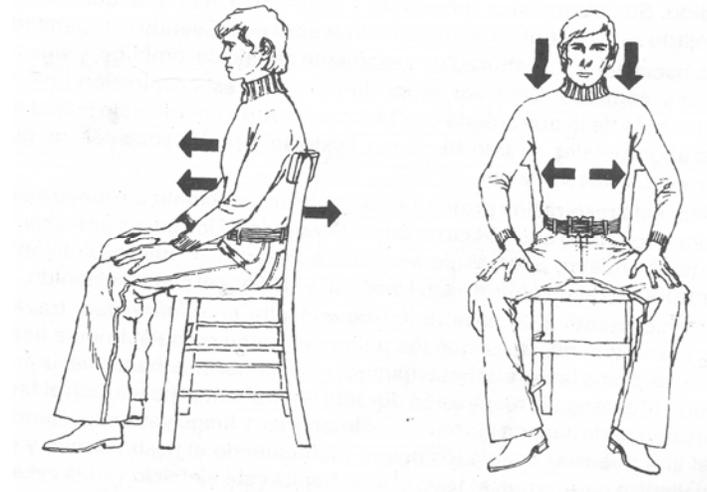
### **1.2.2.2 Ejercicio de respiración completa**

Para realizar este ejercicio se debe sentar sobre una silla con la espalda recta y la cabeza erguida como se muestra en la figura 2. Coloque la mano derecha sobre el vientre y comience a inspirar lentamente por la nariz, sin levantar los hombros, notará como se hincha el vientre, cuando llegue a su máxima capacidad siempre inhalando lentamente comience a inflar la zona media del tórax, notará que las costillas se separan, el vientre se decrecerá para ir llevando el aire a una zona superior y para completar la inspiración debe inflar el pecho reteniendo el aire unos segundos y luego exhalarlo por la boca, produciendo el sonido “fff”, este ejercicio se debe realizarse varias veces.

---

<sup>13</sup>Marcelo Roascio. Lecciones de Canto. Disponible en:  
<http://www.guitarraonline.com.ar/canto2.htm>

**Figura 2. Postura para ejercicio de respiración completa.**



Fuente: Hewitt Graham.  
Manual Cómo Cantar.  
Pág.3.

### **1.2.2.3 Ejercicio de respiración abdominal**

Esta es la respiración que más se utiliza al cantar, por eso es muy importante que se llegue a un gran manejo de esta técnica. Esta respiración se debe practicar estando sentado con la espalda recta luego se coloca la mano sobre el vientre y se inspira lentamente por la nariz. El vientre se hincha, este debe ser el único que se hinche. Luego se exhala por segmentos produciendo el sonido “fff”. Ejemplo: fff-pausa-fff-pausa-fff- pausa, hasta vaciar los pulmones, no se debe inspirar en las pausas. Este ejercicio se debe repetir 5 veces.<sup>14</sup>

---

<sup>14</sup> Marcelo Roascio. Lecciones de Canto. Disponible en:  
<http://www.guitarraonline.com.ar/canto2.htm>

### **1.2.3 Desarrollo de la afinación**

Afinar significa, entre otras cosas “Cantar o ejecutar bien entonado los sonidos”. El oído humano generalmente puede diferenciar entre sonidos afinados y desafinados y dado que la finalidad del canto es agradar al oído humano con la ejecución de de la voz, es necesario que se practique y se realicen ejercicios para mejorar este aspecto en el momento de cantar.

Esta característica del canto es sin duda alguna la que más se busca mejor con la creación del presente manual de canto interactivo proponiendo los siguientes métodos de afinación.

#### **1.2.3.1 Método de afinación 1**

Para lograr desarrollar el oído musical se puede empezar con cualquier instrumento musical, por ejemplo el teclado, ya que genera los tonos musicales de las 7 notas. Se debe tocar un tono largo en el instrumento por ejemplo, un La, luego se debe vocalizar sobre el tono diciendo “laaaaa” hasta que el tono de la voz sea correspondiente al tono ejecutado. Repítalo con distintos tonos que no sean consecutivos entre si, por ejemplo La-Do-Sol.

#### **1.2.3.4 Método de afinación 2**

Se debe jugar con los sonidos basándose en la pregunta y respuesta que hacen los instrumentos musicales, en un solo por lo general. Se debe contestar al instrumento con la voz, a modo de eco, los arpeggios de tres tonos que el instrumento ejecute. Se puede utilizar cualquier vocal o sílabas para seguirlos por ejemplo “La” o “Do” o “e”, etc.

De esta manera se acostumbrará a seguir la música de manera afinada, cosa que todo cantante debe poseer, ya que así es como comienza la composición de una melodía vocal para una canción nueva. También se puede cantar sobre una canción que resulte fácil y escuchar el grado de precisión que se logra, se debe prestar atención a la voz, ya que hay que elegir un instrumento al cual seguir como puede ser la guitarra o el piano.<sup>15</sup>

---

<sup>15</sup> Marcelo Roascio. Lecciones de Canto. Disponible en:  
<http://www.guitarraonline.com.ar/canto2.htm>

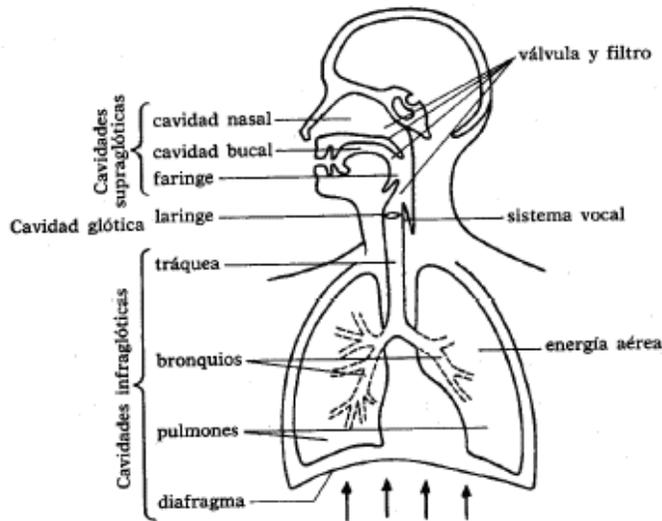


## 2. PROCESO DE PRODUCCIÓN Y CAPTURA DE LA VOZ HUMANA

### 2.1 El aparato fonador

El órgano productor de la voz se compone de tres partes diferentes: las cavidades infraglóticas, la cavidad glótica y las cavidades supraglóticas (ver figura 3); cada una de ellas realiza una función determinada pero todas ellas son necesarias para la producción del habla.<sup>16</sup>

Figura 3. Aparato fonador



Fuente: Mosquera Rivera Ismael.

Sistema de Conversión de Voz a Midi.

Pág.6.

<sup>16</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

La misión de las cavidades infraglólicas es proporcionar la corriente de aire necesaria para producir el sonido, y se componen de diafragma, pulmones, bronquios y tráquea. El diafragma es un músculo que se encuentra situado por debajo de los pulmones y tiene forma de cúpula, cuyo trabajo es controlar, junto con los músculos pectorales el hinchado y vaciado de la cavidad pulmonar produciendo así la respiración.

Cuando el diafragma se contrae, se ensancha la cavidad torácica y se produce la inspiración de aire, y cuando se relaja se reduce la cavidad torácica y se produce la expiración del aire contenido en los pulmones. Los bronquios y la tráquea se encargan de conducir el aire desde los pulmones hacia la laringe, y su cometido es actuar como canales de transmisión del flujo aéreo. El funcionamiento del aparato respiratorio nos sugiere la idea de un compresor.

La cavidad glótica está formada por la laringe, y tiene como característica especial la presencia de las cuerdas vocales, las cuales son las responsables de producir la vibración básica para la generación del sonido. Las cuerdas vocales son en realidad un par de pliegues cubiertos por una membrana mucosa; tienen una longitud aproximada de 3 mm en recién nacidos, y de 9 a 13 mm y 15 a 23 mm en mujeres y hombres adultos respectivamente.

La importancia de la longitud de las cuerdas vocales en la producción del tono de voz fue científicamente analizada recientemente por Sawashima (1983), quien mostró que a mayor longitud, la voz tiene un rango de tonalidades más grave, y a menor longitud, dicho rango está formado por tonalidades más agudas.<sup>17</sup>

Cuando la corriente de aire que proviene de las cavidades infragloticas pasa por la glotis (espacio que queda entre las cuerdas vocales) las hace vibrar, y el tono resultante puede variar en frecuencia e intensidad dependiendo de la presión del aire y de la longitud y masa de las cuerdas vocales; por ejemplo, si un cantante está entonando un A4 (nota La de la cuarta octava), las cuerdas vocales se abren y cierran 440 veces por segundo, es decir, oscilan con una frecuencia de 440 Hz.

Si queremos hacer una buena descripción de las características de la voz debemos definirla en tres dimensiones: frecuencia fundamental, amplitud y espectro; o en términos de acústica: tono, intensidad y timbre. Según esta exposición, en términos de ingeniería podríamos pensar en esta parte del aparato fonador como si se tratase de un oscilador.

---

<sup>17</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

Las cavidades supraglóticas son: faríngea, nasal, bucal y labial. A continuación de la laringe se encuentra la faringe, de donde arranca la raíz de la lengua, y aparece el primer obstáculo móvil: la úvula, que es el apéndice situado al final del velo del paladar o paladar blando. Si la corriente de aire sale exclusivamente por la boca se producen sonidos orales, pero si el velo del paladar está caído, parte del aire será expulsado por la cavidad nasal.<sup>18</sup>

La cavidad nasal carece de elementos móviles, por lo cual tiene una función pasiva en la producción del habla. La lengua es el órgano más móvil de la boca registrando, una elevada actividad durante la producción del habla y está dividida en tres partes: raíz, dorso y ápice.

Recientemente se ha demostrado que dependiendo del perfil que adopta se produce un resonado acústico, por lo que el timbre del sonido será diferente, según la forma sea convexa, cóncava o plana, o esté situada en la parte anterior, central o posterior.

Los dientes son elementos pasivos en la actividad del habla por estar colocados de forma fija en los maxilares; los inferiores tienen movilidad por estar sujetos en la mandíbula inferior, siendo ésta activa en la articulación.<sup>19</sup>

---

<sup>18</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

<sup>19</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

El paladar consta de dos zonas: el paladar duro, situado sobre el hueso palatino, y el paladar blando, mencionado anteriormente, el cual termina en la úvula. Por último, se tiene los labios, los cuales poseen bastante movilidad y permiten la modificación de los sonidos.

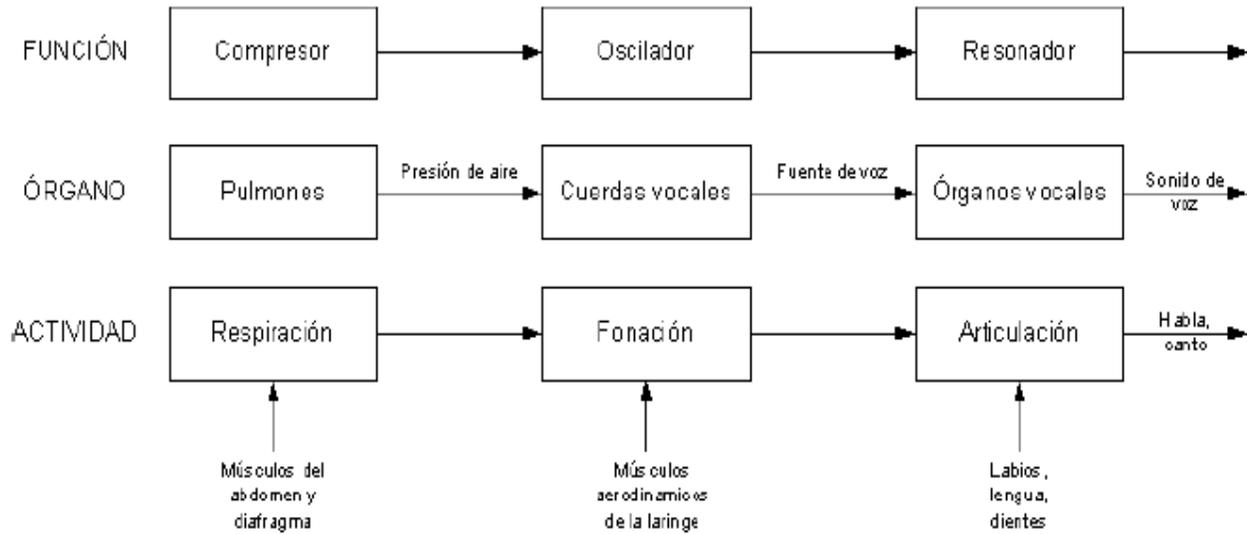
Por tanto, podemos ver que para la producción del habla son necesarias una fuente de energía que expulse aire a presión (aparato respiratorio), un órgano vibratorio (las cuerdas vocales), y una caja de resonancia con sistema de articulación.

A partir de esta información general sobre el funcionamiento del aparato fonador, y una vez marcada la separación en diferentes bloques de las distintas partes que toman parte en la producción del habla, se puede representar el concepto mediante el diagrama de la figura 4.<sup>20</sup>

---

<sup>20</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

**Figura 4. Funcionamiento del aparato fonador.**



Fuente: Mosquera Rivera Ismael

Sistema de Conversión de Voz a Midi

Pág.13.

## 2.2. El aparato auditivo

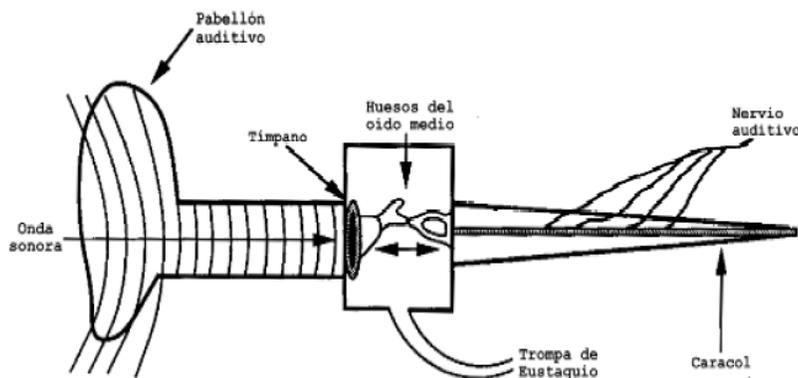
El rango de frecuencias que puede percibir el ser humano se sitúa entre los 20 Hz y los 20000 Hz. En una onda, la intensidad es relativa a la energía (o variación de presión de aire). En general, cuando hay un incremento de intensidad también hay un incremento en la intensidad sonora (atributo perceptual).<sup>21</sup>

<sup>21</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

El oído es más sensible a algunas regiones frecuenciales que a otras. La región más sensible está entre los 2700 y 3200 Hz, es por eso que una onda sinusoidal a 3000 Hz con una determinada intensidad se percibe con una sonoridad mayor que otra a 200 ó 8000 Hz de igual intensidad.

La función del oído es capturar las ondas acústicas y transformarlas en impulsos nerviosos que el cerebro puede interpretar.<sup>22</sup> El oído está formado por tres partes: el oído externo, el oído medio y el oído interno (ver figura 5).

**Figura 5. Aparato auditivo.**



Fuente: Mosquera Rivera Ismael

Sistema de Conversión de Voz a Midi

Pág.15.

---

<sup>22</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

El oído externo está formado por el pabellón auditivo (oreja), y por el conducto auditivo externo.<sup>23</sup> El pabellón auditivo recoge las ondas sonoras y facilita su paso hacia el interior, aplicando una amplificación sobre dichas ondas. El conductor auditivo externo termina en el tímpano, que es una membrana que separa el oído externo del oído medio. El conductor auditivo conduce las ondas hacia el tímpano atenuando los sonidos más agudos que puedan causar daño.

El oído medio está separado del oído externo a través del tímpano, y dispone de una cadena ósea de tres huesecillos: martillo, yunque y estribo.<sup>24</sup> En el oído medio también se encuentra la trompa de Eustaquio, que es un canal que comunica con la faringe. La misión del oído medio es transmitir al oído interno los sonidos que llegan desde el oído externo realizando una adaptación. Cuando la intensidad es pequeña, la cadena ósea se mueve en conjunto produciendo un aumento de la misma; cuando la intensidad es grande, se produce una disminución de la misma con el fin de evitar daños en el oído interno.

---

<sup>23</sup>Dr. Pascual Estrada Estrada. ¿Oye Pero no Entiende? Disponible en: <http://www.drpscuaestrada.com/tema4.htm>

<sup>24</sup> Introducción. Disponible en: <http://www.audicenter.com.ar/introduccion.htm>

En el oído interno se encuentra el caracol, que es el órgano de audición, y cuya función es percibir las frecuencias de las vibraciones sonoras y convertirlas en impulsos nerviosos que transmite al cerebro para su interpretación.<sup>25</sup>

En primer lugar, el sonido es amplificado, para posteriormente equilibrarse; el sonido se transforma en su representación a través del caracol, del cual sale el conjunto de señales nerviosas que llega al cerebro.

### 2.3 El sonido

El sonido se define como la sensación producida en el órgano del oído por el movimiento vibratorio de los cuerpos, transmitido por un medio elástico como el aire. Entonces podemos deducir que el origen del sonido es la vibración de los cuerpos.<sup>26</sup> Se dice que la vibración es periódica cuando se repite a intervalos regulares.

La representación matemática que describe el movimiento vibratorio que se produce cuando golpeamos un diapasón puede expresarse mediante la función coseno:  $x(t) = A\cos(\omega t + \varphi) = A\cos(2\pi f t + \varphi)$ <sup>27</sup>

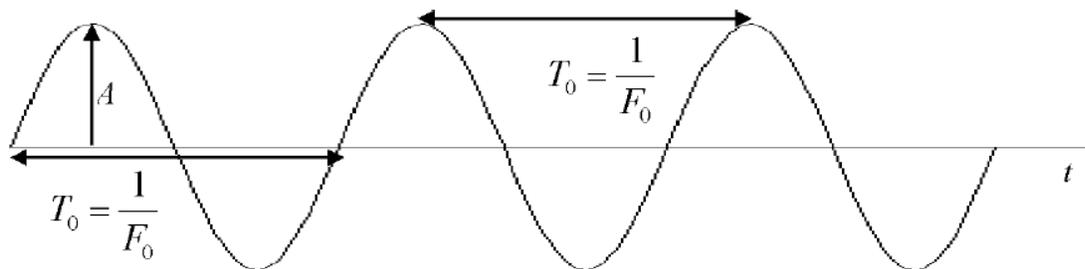
---

<sup>25</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

<sup>26</sup> Maikel Galeon. El Disco Compacto. Disponible en: <http://maikel.galeon.com/cd/ElDiscoCompacto.html>

<sup>27</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

Figura 6. Gráfica de la función  $x(t) = A\cos(\omega t + \phi)$



Fuente: Mosquera Rivera Ismael

Sistema de Conversión de Voz a Midi.

Pág.11. A: amplitud, T: periodo, F: frecuencia

Para generar esta función mediante un ordenador se debe pasar la variable continua  $t$  a la forma discreta  $Tn$ , con  $n = 0, 1, \dots$  y donde  $T$  representa el período de muestreo o intervalo entre dos muestras, siendo la frecuencia de muestreo  $F_s = 1/T$ . La diferencia entre la forma continua y la discreta es que la segunda no está definida en todo el eje temporal, sino que sólo se define en instantes concretos equiespaciados en el tiempo.

Cualquier sonido complejo puede ser formado como la suma de armónicos simples de distinta frecuencia, amplitud y fase; que corresponden a los parciales de la señal sonora.

Así pues, para analizar una señal de audio se debe descomponerla en el conjunto de armónicos simples que la componen, ya que sus intensidades y frecuencias proporcionan información sobre las características del sonido original.

La mayoría de sonidos producidos por la voz son cuasi periódicos, por tanto, podemos aplicar este concepto para su análisis. Por otra parte, el sonido aperiódico o ruido es una onda que no presenta ninguna periodicidad y al descomponerla se observa que la energía está dispersa entre sus armónicos y carece de estabilidad.<sup>28</sup>

## **2.4 Características musicales del sonido**

En el sonido distinguimos diversos elementos, como la intensidad o fuerza con que se produce este sonido; la altura que nos hace considerarlo como agudo, medio o grave; el timbre, que es aquella cualidad del sonido gracias a la cual sabemos que lo que escuchamos está producido por uno u otro instrumento o voz; y la duración que nos permite apreciar el tiempo que el sonido está en nuestro oído.

### **2.4.1 Frecuencia**

Cada sonido se caracteriza por su velocidad específica de vibración, que impresiona de manera peculiar al sentido auditivo. Esta propiedad recibe el nombre de tono.<sup>29</sup>

---

<sup>28</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

<sup>29</sup> La Caja de Música. Las Características Principales del Sonido. Disponible en: [http://www.xtec.es/centres/a8019411/caixa/musica\\_es.htm](http://www.xtec.es/centres/a8019411/caixa/musica_es.htm)

La duración de las vibraciones de instrumentos de un mismo tipo es proporcional a sus dimensiones lineales. Los sonidos de mayor o menor frecuencia se denominan respectivamente, agudos o graves; términos relativos, ya que entre los tonos diferentes uno de ellos será siempre más agudo que el otro y a la inversa.<sup>30</sup>

### **2.4.2 Amplitud**

Es la cualidad que permite distinguir entre sonidos fuertes o débiles. Además de la amplitud en la percepción de la intensidad, influye la distancia a que se encuentra situado el foco sonoro del oyente y la capacidad auditiva de este.

### **2.4.3 Forma de onda.**

Si el tono permite diferenciar unos sonidos de otros por su frecuencia, y la intensidad los sonidos fuertes de los débiles, el timbre completa las posibilidades de variedades del arte musical desde el punto de vista acústico, porque es la cualidad que permite distinguir los sonidos producidos por los diferentes instrumentos.<sup>31</sup>

---

<sup>30</sup>La Caja de Música. Las Características Principales del Sonido. Disponible en:  
[http://www.xtec.es/centres/a8019411/caixa/musica\\_es.htm](http://www.xtec.es/centres/a8019411/caixa/musica_es.htm)

<sup>31</sup> La Caja de Música. Las Características Principales del Sonido. Disponible en:  
[http://www.xtec.es/centres/a8019411/caixa/musica\\_es.htm](http://www.xtec.es/centres/a8019411/caixa/musica_es.htm)

## 2.5 Archivos de audio

Los archivos de audio permiten analizar en función del tiempo el nivel de sonido así como su composición espectral (frecuencias). En un archivo digital de sonido se almacenan, usando números enteros, la amplitud (nivel de presión sonora) de uno o más canales de audio en función del tiempo. Las muestras de amplitud se toman a una frecuencia fija FS y se emplean también un número fijo de Bytes NB para cada uno de los canales NC.

Cuanto mayores sean estos números mayor será la calidad del sonido, pero también crecerá de forma lineal el tamaño de los archivos. En la tabla I se muestra la tabla con los valores más típicos de estos parámetros y el número de octetos necesarios para registrar (sin compresión) 1 segundo de sonido.<sup>32</sup>

**Tabla I. Parámetros típicos de la digitalización de sonido.**

calidad	FS [Hz]	NB	NC	tamaño de 1 segundo
teléfono	11025	8	1	11 mil octetos
radio	22050	8	1 (mono)	22 mil octetos
CD	44100	16	2 (estéreo)	172 mil octetos

Fuente: Rego Alberto Ramil

Instrucciones generales para el análisis de archivos multimedia

Pág.14.

---

<sup>32</sup> Alberto Ramil Rego. Instrucciones Generales Para el Análisis de Archivos Multimedia Disponible en: [http://www.ii.udc.es/lail/em/laboratorio\\_virtual/DOC/analisis.pdf](http://www.ii.udc.es/lail/em/laboratorio_virtual/DOC/analisis.pdf)

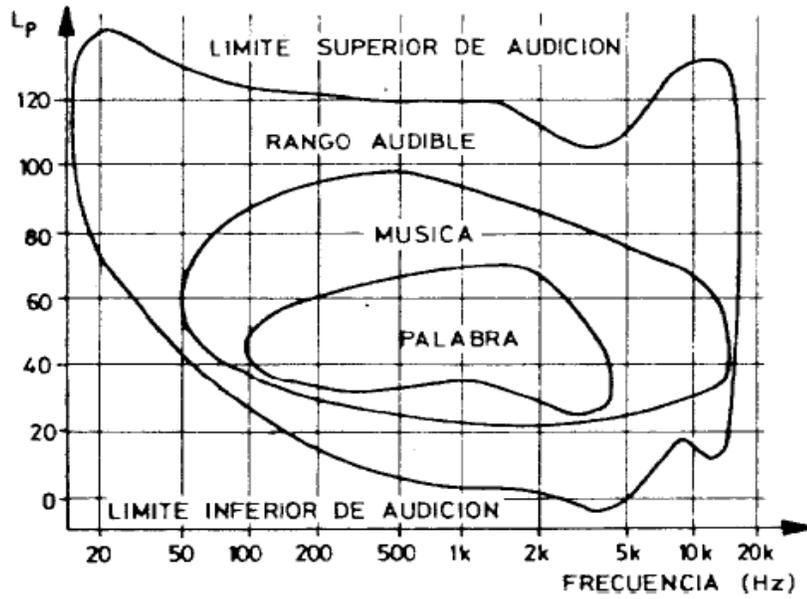
Los valores de los parámetros se eligen de acuerdo con la aplicación particular y con las capacidades del oído humano, que son enormes. Debe mencionarse que la respuesta del oído, tanto a la intensidad como a la frecuencia no es lineal sino más bien logarítmica. Así, los oídos humanos más sensibles son capaces de detectar cambios de intensidad de 0.3 dB y cambios en la frecuencia de vibración (tono) que correspondan al 0,03% de la frecuencia original, en el rango comprendido entre 500 y 8.000 hertzios. Esta es la razón por la que en acústica se utilizan tan frecuentemente escalas logarítmicas.

Si nos fijamos en los parámetros de grabación de un CD vemos que el ancho de banda  $FS/2=22050$  es suficiente para la mayoría de los instrumentos musicales que producen muy pocos sonidos por encima de 10 KHZ , y que los 16 bites de amplitud equivalentes a 96 dB de intensidad son suficientes para cubrir las diferencias de intensidad de una orquesta entre el “pianissimo ”=20 dB y el “fortissimo” = 100 dB.<sup>33</sup>

---

<sup>33</sup> Alberto Ramil Rego. Instrucciones Generales Para el Análisis de Archivos Multimedia Disponible en: [http://www.ii.udc.es/lail/em/laboratorio\\_virtual/DOC/analisis.pdf](http://www.ii.udc.es/lail/em/laboratorio_virtual/DOC/analisis.pdf)

Figura 7. Áreas de audición.



Fuente: Rego Alberto Ramil

Instrucciones generales para el análisis de archivos multimedia

Pág.15.



## 3. ANÁLISIS ESPECTRAL DE LA VOZ HUMANA

### 3.1 Análisis espectral

En el siglo XIX, científicos y músicos ya conocían que muchos sonidos musicales estaban caracterizados por vibraciones armónicas alrededor de un tono fundamental, pero no disponían de la tecnología necesaria para analizar esos armónicos de una forma sistemática. Isaac Newton empleó el término “espectro”, para describir las bandas de color que muestran las diferentes frecuencias cuando se descompone la luz blanca al pasar a través de un prisma de cristal.

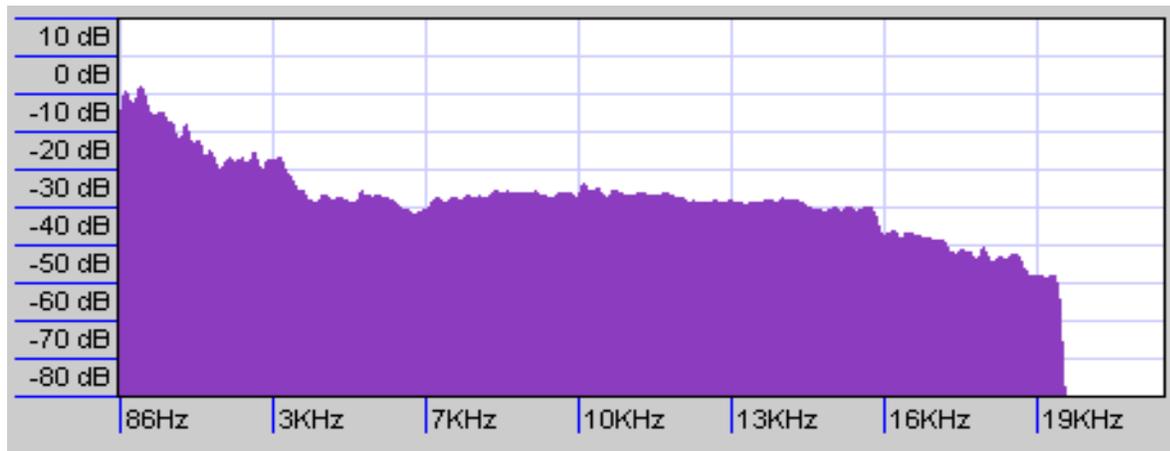
Tal como una imagen puede describirse como una mezcla de colores, un sonido puede ser descrito como una combinación de vibraciones acústicas elementales. Una forma de diseccionar un sonido es considerar la contribución de varios componentes, cada uno correspondiente a una cierta variación en la presión del aire. Medir la distribución entre estos componentes se denomina análisis espectral.<sup>34</sup>

---

<sup>34</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

El análisis espectral revela las características de frecuencia y energía en tonos vocales e instrumentales, lo cual lo hace valioso en la detección de la tonalidad, factor imprescindible en el contexto de este proyecto. La gráfica del espectro de una señal sonora (ver figura 8) se representa en un plano bidimensional cuyos ejes son la amplitud y la frecuencia. De entre los métodos existentes empleados para la realización de un análisis espectral, el más ampliamente utilizado es el análisis de Fourier..<sup>35</sup>

**Figura 8. Gráfica de espectro de señal sonora**



Fuente: Programa Edición y grabación de Audio Audacity.

Dado el amplio campo de aplicaciones de la transformada de Fourier y la gran cantidad de proyectos en los cuales se utiliza como herramienta de análisis de señales se ha designado a la transformada de Fourier como la herramienta de análisis espectral a utilizar en la presente aplicación.

---

<sup>35</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

### 3.1.1 Transformada de Fourier

En 1822 el ingeniero francés Jean-Baptiste Joseph, Baron de Fourier publicó su tesis sobre teoría analítica del calor; en ella describía como complicadas vibraciones pueden ser analizadas como suma de varias señales simples; en particular, demostró que cualquier función periódica puede ser representada como una suma infinita de senos y cosenos.

La transformada de Fourier es una herramienta matemática que nos permite obtener el espectro de una señal; la formulación de la integral de Fourier es:<sup>36</sup>

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt$$

Lo cual indica que la FT (Fourier Transform) de una determinada frecuencia  $f$  es la integral de la multiplicación de la señal de entrada  $x(t)$  por el tono puro  $e^{-j\omega t}$ . El espectro de frecuencias entre 0 Hz hasta infinito con su correspondiente imagen especular en el eje negativo de frecuencias es  $X(f)$ , que es la notación empleada para representar la FT.

Los sistemas digitales tratan con señales discretas, por esa razón es necesaria una representación en tiempo discreto de la FT (señal en tiempo continuo); dicha solución viene dada por la DTFT (Discrete Time Fourier Transform).

---

<sup>36</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

### 3.1.2 Transformada de Fourier discreta (DTF) Y transformada de Fourier rápida (FFT).

Una señal discreta es una señal analógica que ha sido muestreada. Después del muestreo, la señal pasa a través de un filtro anti-aliasing, con el fin de atenuar (eliminando al máximo) las frecuencias que están por encima de la mitad de la velocidad de muestreo (tasa de Nyquist), produciendo una nueva señal discreta. La notación para las señales en tiempo discreto se representa con el índice temporal entre corchetes:  $x[n]$ .<sup>37</sup>

Así pues, si muestreamos una señal  $x(t)$  con una frecuencia de muestreo  $F_s$  obtenemos la señal discreta  $x[n]$ . El periodo de muestreo  $T$  es igual a  $1/F_s$  por lo que la señal resultante es  $x(nT)$ . La transformada de Fourier en tiempo discreto DTFT puede denotarse como  $X(e^{j\theta})$ , donde  $\theta$  es la fase que toma valores positivos y negativos.

La DTFT revela el espectro continuo de frecuencias de la señal  $x[n]$ , que sólo está definida para valores de la variable entera  $n$ , pero que está compuesta por un número infinito de muestras. Si tenemos en cuenta que una exponencial compleja es idéntica a otra que difiere de ella en frecuencia en múltiplos de  $2\pi$ , entonces podemos representarla hasta el infinito en intervalos de  $2\pi$  correspondientes a múltiplos de la velocidad de muestreo  $F_s$ . La DTFT puede ser representada como:

$$X(e^{j\theta_0}) = X(e^{j(\theta_0 + 2\pi k)})$$

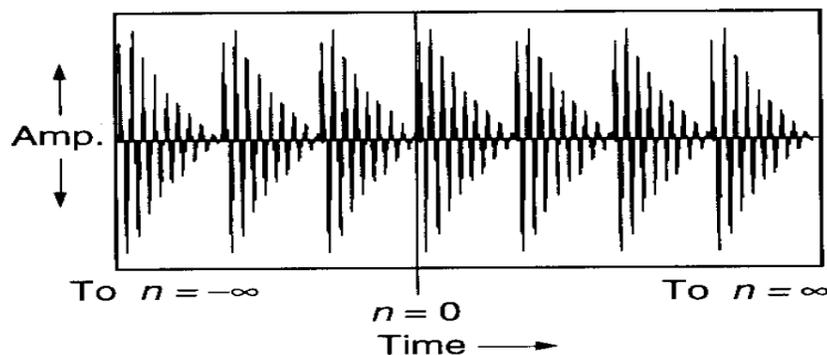
---

<sup>37</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

En donde  $k$  es un múltiplo arbitrario. En la práctica sólo es necesario considerar los valores de la DTFT en un período de  $2\pi$ , y se asume que la DTFT  $X(ej\omega)$  entre  $-\pi$  y  $\pi$  está libre de aliasing, pero  $X(ej\omega)$  es un espectro infinito y una computadora puede calcular solamente cantidades discretas, por lo que necesitamos una técnica de análisis que trabaje con espectros discretos de señales discretas.<sup>38</sup>

Esta técnica es la transformada de Fourier discreta: DFT. Podemos construir la señal  $\tilde{x}[n]$  replicando la señal  $x[n]$  hasta el infinito, y la DFT de  $\tilde{x}[n]$  se compone de muestras de la DTFT de  $x[n]$ . Así como muestrear  $x(t)$  en el dominio temporal resulta en replicar en el dominio frecuencial; replicar  $x[n]$  en el dominio temporal resulta en muestrear en el dominio frecuencial. En las figuras 9 y 10 siguientes se muestra este concepto de forma gráfica.<sup>39</sup>

Figura 9. Gráfica de la señal  $\tilde{x}[n]$

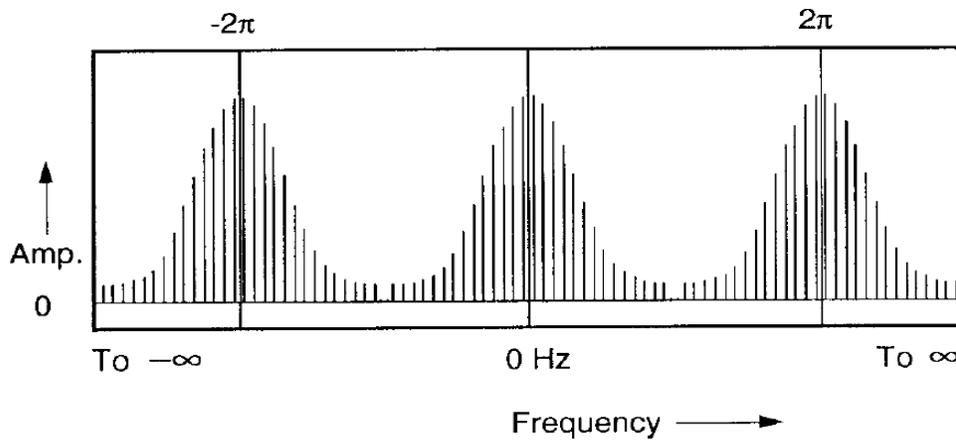


Fuente: Mosquera Rivera Ismael  
 Sistema de Conversión de Voz a Midi  
 Pág. 17.

---

<sup>38</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

Figura 10. DFT de la señal muestreada  $\tilde{x}[n]$



Fuente: Mosquera Rivera Ismael

Sistema de Conversión de Voz a Midi

Pág.17.

En la figura 9 se muestra que la señal  $\tilde{x}[n]$  consiste en réplicas hasta el infinito de la señal  $x[n]$ . En la figura 10 vemos una gráfica de la DFT de  $\tilde{x}[n]$ , que resulta en un espectro discreto en frecuencia.

La FFT (Fast Fourier Transform) no es más que una implementación eficiente de la DFT para conseguir un rendimiento computacional aceptable. La DFT tiene un orden de  $N^2$  multiplicaciones complejas, mientras que el orden de la FFT es de  $N \log(N)$  operaciones, lo cual es una importante optimización en el rendimiento, si tenemos en cuenta además la gran cantidad de datos que intervienen en el análisis de audio.<sup>40</sup>

---

<sup>39</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

<sup>40</sup> Ismael Mosquera Rivera. Sistema de Conversión de Voz a Midi. Disponible en: <http://www.create.ucsb.edu/~xavier/pfc/imosquera.pdf>

### **3.2 Propiedades de la voz digitalizada.**

La señal digitalizada transmite por sí misma poca información y el propósito del procesamiento de voz es aplicar transformaciones apropiadas a este material crudo, para derivar rasgos representativos de la señal bajo análisis.

Estos rasgos varían más lento respecto a las variaciones rápidas de amplitud de la señal original, y se relacionan de alguna manera con la secuencia de fonemas que contiene la información lingüística. Otros rasgos pueden relacionarse con características más cualitativas como la expresión, la altura de la voz, la identidad del hablante, etc. La tasa de cambio de estos rasgos es de solo 4 a 5 por segundo, si bien esto no es constante y pueden ocurrir cambios rápidos entre un estado estacionario y otro.

Al no tener conocimiento previo de la estadística de la señal de voz, las muestras se consideran como provenientes de una fuente aleatoria de distribución desconocida. El objetivo de los algoritmos de procesamiento es hacer estimaciones de los parámetros estadísticos locales. Una estimación precisa puede hacerse únicamente en el límite de un número infinito de muestras de una distribución estacionaria.<sup>41</sup>

---

<sup>41</sup>Ernesto López. Búsqueda de Música por tarareo. Disponible en:  
<http://iie.fing.edu.uy/investigacion/grupos/gmm/proyectos/tararira/descargas/tararira.pdf>

Pero para que la señal de voz contenga información, la estadística de la misma debe cambiar frecuentemente, siendo la distribución de muestras de la señal de voz fuertemente no estacionaria. Sin embargo, se puede asumir que en tramos cortos de tiempo, pequeñas secuencias de muestras pueden ser localmente estacionarias. De esta forma, los algoritmos pueden trabajar sobre pequeños bloques de muestras consecutivos. Desafortunadamente, esto implica que la precisión de las estimaciones será limitada.

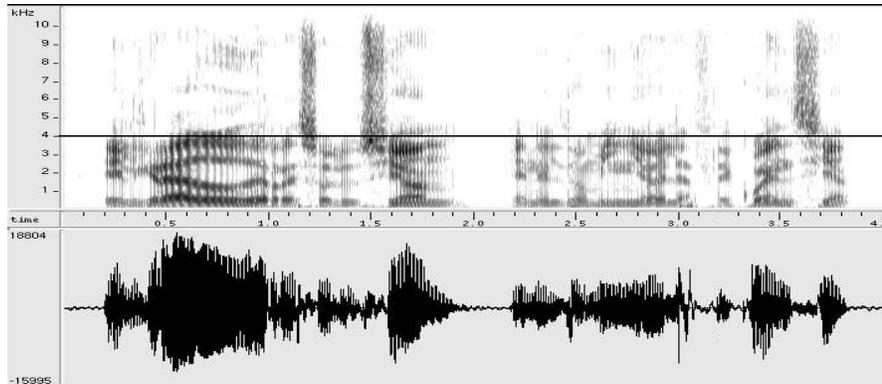
### **3.2.1. Ancho de banda y frecuencia de muestreo**

La figura 11 muestra en un espectrograma la distribución de energía de una señal de voz. A efectos ilustrativos se trazó una línea a 4kHz. Se puede apreciar que la mayor parte de la energía se concentra por debajo de este ancho de banda, a pesar de que hay componentes de frecuencia más altos (sonidos sordos). Si bien depende de la aplicación, en general una frecuencia de muestreo ( $f_s$ ) de 8 kHz es adecuada para el procesamiento de voz. En aplicaciones en las que se pretenda tratarla señal de voz con más precisión (por ejemplo en música) serán necesarias  $f_s$  más altas, tales como 22050 Hz o 44100 Hz.<sup>42</sup>

---

<sup>42</sup> Ernesto López. Búsqueda de Música por tarareo. Disponible en:  
<http://iie.fing.edu.uy/investigacion/grupos/gmm/proyectos/tararira/descargas/tararira.pdf>

**Figura 11. Espectrograma de señal de voz.**



Fuente: López Ernesto, Rocamora Martín, Sosa Gonzalo  
Búsqueda de Música por tarareo.  
Pág.9.

### **3.2.2 Largo del bloque de análisis**

La elección del largo del bloque de análisis es un tema crucial. Si es muy corto, el algoritmo puede no tener suficientes muestras para funcionar adecuadamente, y si es muy largo los parámetros estadísticos cambian dentro del bloque dando lugar a resultados que son un promedio más a largo plazo.

En la práctica bloques de 20-30ms (160 a 240 muestras a  $f_s = 8\text{kHz}$ ), son adecuados para la mayoría de las aplicaciones.<sup>43</sup>

---

<sup>43</sup> Ernesto López. Búsqueda de Música por tarareo  
<http://iie.fing.edu.uy/investigacion/grupos/gmm/proyectos/tararira/descargas/tararira.pdf>

### 3.2.3 Enventanado

Seleccionar un bloque de muestras es enventanar la señal, y este proceso es de mucha relevancia. Una ventana rectangular tiene una forma de *sinc* en frecuencia con lóbulos secundarios amplios. La convolución en frecuencia de este espectro con el de la señal produce distorsión considerable. Si el algoritmo de estimación involucra análisis en el dominio de la frecuencia (como DFT), deberán ser usadas ventanas más suaves en el dominio del tiempo (como Hamming o Hanning).

Si se pretenden detectar cambios rápidos en la señal de voz, es común procesar la señal en bloques que se solapan en el tiempo. Por ejemplo, para análisis espectral preciso, cada DFT puede ser hecha sobre bloques de 32ms (256 muestras a 8kHz), pero desplazados solo 5 ms. Los rasgos espectrales de bloques adyacentes serán similares durante tramos estacionarios de la señal, mientras que podrán detectarse cambios temporales de hasta 5 ms.<sup>44</sup>

### 3.2.4 Resolución espectral

La resolución que se obtiene en el dominio de la frecuencia está dada por el largo del bloque de análisis  $N$ , es decir, la cantidad de puntos de la señal usados para calcular la DFT, siendo esta de  $f_s/N$ . Esto implica el compromiso existente entre resolución temporal y resolución espectral. Para  $f_s = 8\text{kHz}$  y 200 muestras, la resolución es de 40Hz. Esto indica que rasgos espectrales separados menos de 40Hz no se pueden resolver.

---

<sup>44</sup> Ernesto López. Búsqueda de Música por tarareo  
<http://iie.fing.edu.uy/investigacion/grupos/gmm/proyectos/tararira/descargas/tararira.pdf>

Para la señal de voz, el detalle espectral más no corresponde al tono de una voz grave masculina, y este es raras veces menor a 60Hz. En general transformadas de 200 a 250 muestras son adecuadas para análisis de voz (a  $f_s = 8\text{kHz}$ ).<sup>45</sup>

---

<sup>45</sup> Ernesto López. Búsqueda de Música por tarareo  
<http://iie.fing.edu.uy/investigacion/grupos/gmm/proyectos/tararira/descargas/tararira.pdf>



## 4. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE MANUAL DE CANTO INTERACTIVO, BASADO EN LA TÉCNICA DE CANTO ALEMAN

### 4.1 Análisis y diseño

Se denomina análisis de sistemas a una de las etapas de construcción de un sistema informático, que consiste en revelar la información actual y proponer los rasgos generales de la solución futura <sup>46</sup> y se denomina *diseño de sistemas* al arte de definir la arquitectura de hardware y software, componentes, módulos y datos de un sistema de cómputo para satisfacer ciertos requerimientos. Es la etapa posterior al análisis de sistemas.<sup>47</sup>

#### 4.1.1 Requerimientos funcionales

Un requerimiento funcional define el comportamiento interno del software: cálculos, detalles técnicos, manipulación de datos y otras funcionalidades específicas que muestran cómo los casos de uso serán llevados a la práctica.<sup>48</sup>

---

<sup>46</sup> Colaboradores de Wikipedia. Análisis de Sistemas. Disponible en: [http://es.wikipedia.org/wiki/Análisis\\_de\\_sistemas](http://es.wikipedia.org/wiki/Análisis_de_sistemas)

<sup>47</sup> Colaboradores de Wikipedia. Diseño de Sistemas. Disponible en: [http://es.wikipedia.org/wiki/Diseño\\_de\\_sistemas](http://es.wikipedia.org/wiki/Diseño_de_sistemas)

<sup>48</sup> Colaboradores de Wikipedia. Requerimiento Funcional. Disponible en: [http://es.wikipedia.org/wiki/Requerimiento\\_funcional](http://es.wikipedia.org/wiki/Requerimiento_funcional)

Captura de señal de audio: El sistema de permitir la captura de los ejercicios de canto realizados por el alumno.

- Capacidad para cargar y guardar archivos de audio: El sistema permitirá cargar las lecciones grabadas por el maestro y guardar temporalmente el ejercicio grabado por el alumno.
- Reproducción de audio: La aplicación reproducirá el audio de las muestras grabadas previamente por el maestro así como los ejercicios grabados por el alumno.
- Graficar las señales de audio originales: El sistema graficará la amplitud de las señales grabadas tanto por el alumno como por el maestro en el dominio del tiempo.
- Transformación de las señales de audio: El sistema aplicará la transformada de Fourier a las señales de audio del maestro y del alumno.
- Segmentación: Se fragmentará cada una de las señales luego de ser transformadas para permitir una mejor comparación de las frecuencias en el dominio del tiempo.
- Interfaz gráfica de usuario (GUI). La herramienta contará con una interfaz gráfica de usuario para permitir la interacción con el sistema.
- Visualización de los resultados. El sistema realizará una representación de la forma de onda de las señales evaluadas reflejando claramente la frecuencia de cada una de las señales en el dominio del tiempo.

### 4.1.2 Requerimientos no funcionales

Un requerimiento no funcional es, en la ingeniería de sistemas y la ingeniería de software, un requerimiento que especifica criterios que pueden usarse para juzgar la operación de un sistema en lugar de sus comportamientos específicos, ya que éstos corresponden a los requerimientos funcionales.<sup>49</sup>

- Documentación: El sistema estará claramente documentado por medio del presente documento y el manual de usuario.
- Eficiencia: El sistema será eficiente realizando los procedimientos de captura y comparación con un 90% de precisión.
- Mantenimiento: El sistema será fácil de mantener ya que se brindará el código fuente y el manual técnico para futuras readecuaciones y extensiones.
- Rendimiento: El sistema tendrá un rendimiento rápido para el manejo de audio y transparencia para el usuario.
- Precio: El precio de la aplicación será realmente económico ya que forma parte de un proyecto de investigación con fines no lucrativos.
- Calidad: El sistema brindara una herramienta de alta calidad para el aprendizaje del canto.

---

<sup>49</sup> Colaboradores de Wikipedia.Requerimientos no Funcionales. Disponible en: [http://es.wikipedia.org/wiki/Requerimiento\\_no\\_funcional](http://es.wikipedia.org/wiki/Requerimiento_no_funcional)

- Necesidad de recursos: Los recursos necesarios para el uso del sistema son mínimos y comunes en el mercado.
  - Estabilidad: Aún cuando el análisis de audio es complejo el sistema se mantendrá estable provocando la menor cantidad de problemas al usuario.
  - La extracción de las características necesarias para la conversión se llevará a cabo utilizando técnicas de análisis espectral.
  - Los parámetros empleados para la comparación serán la energía y la frecuencia fundamental de la señal (observando sus variaciones en el tiempo).
- El lenguaje de programación utilizado será el lenguaje m de Matlab.
  - La aplicación trabajará con archivos de audio en formato wave.
  - La captura de la señal se hará con una velocidad de muestreo de 44100 Hz y una resolución de 16 bits.
  - La interfaz de usuario se construirá usando GUIDE (*graphical user interface development environment*) de MATLAB.
  - Para la visualización de los datos, se usará el API GUIDE (*graphical user interface development environment*) de MATLAB.
  - La interfaz ofrecerá al usuario una forma cómoda e intuitiva para la utilización del sistema.

## **4.1.3 Herramientas utilizadas**

### **4.1.3.1 Cool Edit Pro 2**

Es una aplicación en forma de estudio de sonido destinado para la edición de audio digital de Adobe Systems Incorporated que permite tanto un entorno de edición mezclado de ondas multipista no-destructivo como uno destructivo.<sup>50</sup>

### **4.1.3.2 Matlab 7.6.0 R2008a**

Matlab (abreviatura de *MATrix LABoratory*, "laboratorio de matrices") es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M).

Entre sus prestaciones básicas se hallan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware.<sup>51</sup>

---

<sup>50</sup> Colaboradores de Wikipedia.Adobe\_Audition. Disponible en:  
[http://es.wikipedia.org/wiki/Adobe\\_Audition](http://es.wikipedia.org/wiki/Adobe_Audition)

<sup>51</sup> Colaboradores de Wikipedia.Matlab. Dispoible en:  
<http://es.wikipedia.org/wiki/MATLAB>

El paquete MATLAB dispone de dos herramientas adicionales que expanden sus prestaciones, a saber, Simulink (plataforma de simulación multidominio) y GUIDE (editor de interfaces de usuario - GUI). Además, se pueden ampliar las capacidades de MATLAB con las *cajas de herramientas (toolboxes)*; y las de Simulink con los *paquetes de bloques (blocksets)*.<sup>52</sup>

## **4.2 Arquitectura**

La arquitectura de sistemas establece los fundamentos para que analistas, diseñadores, programadores, ect. Para que estos trabajen en una línea común que permita alcanzar los objetivos del sistema de información, cubriendo todas las necesidades.<sup>53</sup>

### **4.2.1 Micrófono**

El micrófono permite convertir las señales acústicas a señales análogas por medio de las pastillas censorsas con las que cuentas para luego ser transportadas por un cable hacia la tarjeta de audio.

### **4.2.2 Tarjeta de audio**

La tarjeta de audio es utilizada para ingresar a la computadora las señales análogas generadas por el micrófono y convertirles a digitales para ser manejadas por el sistema.

---

<sup>52</sup> Colaboradores de Wikipedia. Matlab. Disponible en: <http://es.wikipedia.org/wiki/MATLAB>

<sup>53</sup> Colaboradores de Wikipedia. Arquitectura de Software. Disponible en; [http://es.wikipedia.org/wiki/Arquitectura\\_de\\_software](http://es.wikipedia.org/wiki/Arquitectura_de_software)

La tarjeta de audio tiene un relación de doble vía con el usuario, ya que también es la encargada de enviar hacia los audífonos o parlantes las señales reproducidas desde la aplicación.

#### **4.2.3 Interfaz gráfica de usuario**

La interfaz gráfica de usuario es utilizada para que el usuario tenga interacción con la aplicación, la cual permite que el usuario brinde información a la aplicación para ser analizada para luego brindar al usuario la retroalimentación respectiva (resultados).

#### **4.2.4 Funciones de I/O de Matlab**

Matlab y sus funciones de manejo de flujos de entrada y salida permiten el manejo de las señales de entrada y de salida utilizadas por la aplicación.

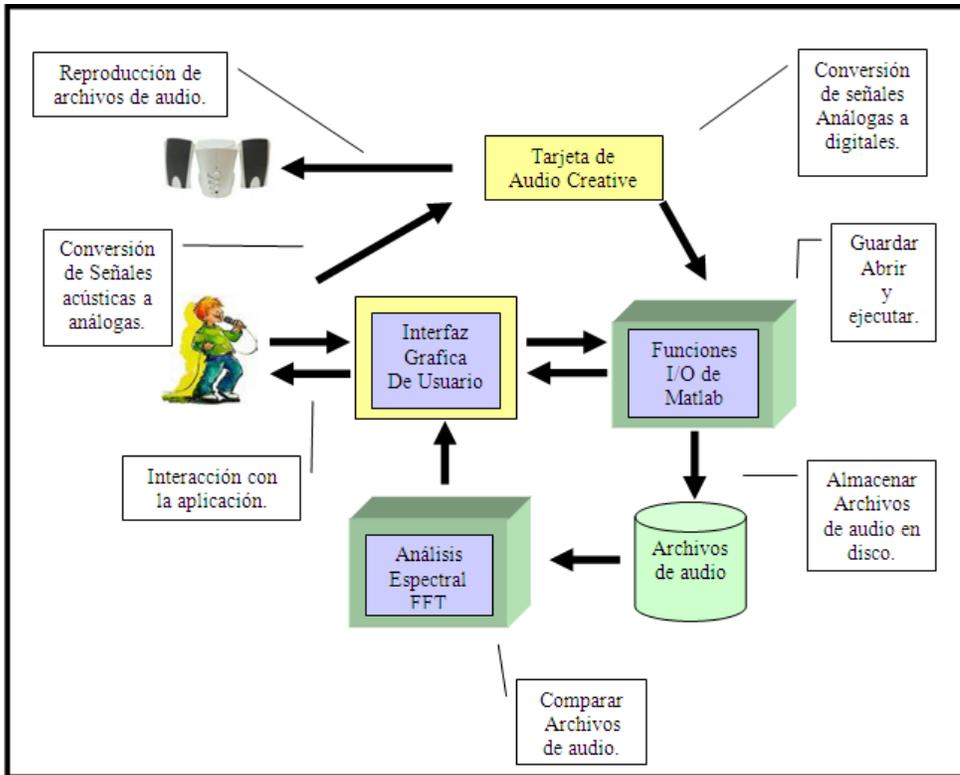
#### **4.2.5 Archivos de audio**

Toda las señales manejadas en la aplicación son guardadas en archivos de audio con formato wave para el mejor manejo en el computador.

#### **4.2.6 Analisis espectral con la función FFT de Matlab**

Matlab y las funciones FFT para realizar la transformada de Fourier de una señal de entrada es una de las bondades de matlab que lo hacen un lenguaje robusto en el procesamiento rápido de grandes cantidades de datos.

Figura 12. Arquitectura del Sistema



### 4.3 Paradigma de programación

Un paradigma de programación representa un enfoque particular o filosofía para la construcción del software. No es mejor uno que otro sino que cada uno tiene ventajas y desventajas. También hay situaciones donde un paradigma resulta más apropiado que otro.<sup>54</sup>

---

<sup>54</sup> Colaboradores de Wikipedia. Paradigma de Programación. Disponible en: [http://es.wikipedia.org/wiki/Paradigma\\_de\\_programaci%C3%B3n](http://es.wikipedia.org/wiki/Paradigma_de_programaci%C3%B3n)

### **4.3.1 Programación orientada a eventos**

La programación dirigida por eventos es un paradigma de programación en el que tanto la estructura como la ejecución de los programas van determinados por los sucesos que ocurran en el sistema o que ellos mismos provoquen.

En la programación dirigida por eventos, al comenzar la ejecución del programa se llevarán a cabo las inicializaciones y demás código inicial y a continuación el programa quedará bloqueado hasta que se produzca algún evento. Cuando alguno de los eventos esperados por el programa tenga lugar, el programa pasará a ejecutar el código del correspondiente administrador de evento.<sup>55</sup>

## **4.4 Implementación de Manual de Canto Interactivo basado en la Técnica de Canto Alemana**

### **4.4.1 Grabación de lecciones**

Las lecciones son los ejercicios grabados por un profesional del canto. En estas lecciones se ponen en práctica las técnicas básicas de afinación y tiempo en el aprendizaje del canto.

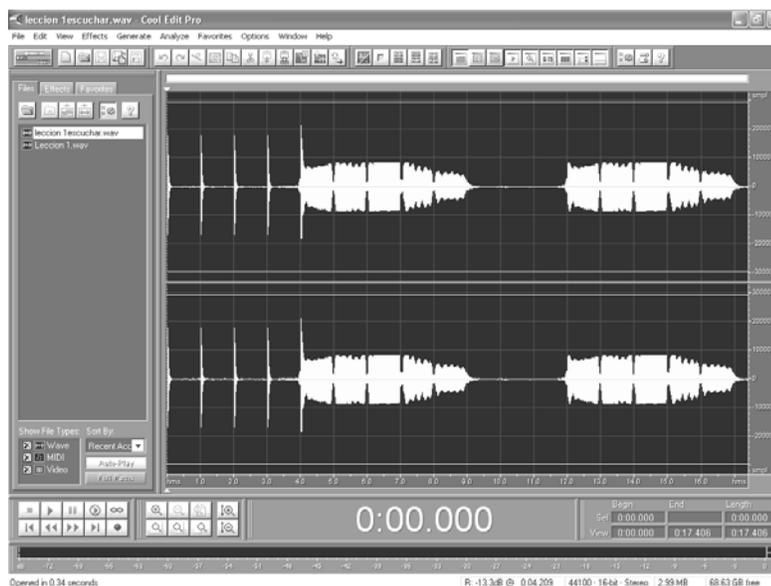
---

<sup>55</sup> Colaboradores de Wikipedia. Programación Dirigida por Eventos. Disponible en: [http://es.wikipedia.org/wiki/Programaci%C3%B3n\\_dirigida\\_por\\_eventos](http://es.wikipedia.org/wiki/Programaci%C3%B3n_dirigida_por_eventos)

Para fines de este proyecto se realizó la grabación de tres lecciones utilizando la herramienta cool edit pro 2, esta herramienta cuenta con un metrónomo que sirve como instrumento para llevar el compas de cada uno de los ejercicios y así facilitar al alumno y al maestro la realización del ejercicio y coincidir en tiempo y en tono.

A continuación se observa la gráfica de tiempo versus amplitud generado en el momento de la grabación de la lección 1 (ver figura 13). En dicha imagen se observan claramente los primeros 4 tiempos marcados por el metrónomo y luego las dos iteraciones del ejercicio propuesto.

**Figura 13. Gráfica de señal de audio de tiempo versus amplitud.**



**Fuente: Programa de edición de audio Cool Edit Pro**

Las características de las lecciones grabadas por el maestro son las siguientes:

- Frecuencia de muestreo ( $F_s$ )=44100
- Número de bits por muestra=16 bits
- Número de Canales=2
- Formato=WAV

A continuación se presentan las líneas de código (Lenguaje m de Matlab) encargadas de realizar la carga de la lección seleccionada.

CAR\_LEC Función encargada de cargar las lecciones

```
function car_lec(t,g,b,leccion)
    %t maneja el objeto texto que muestra el nombre de la lección
    %seleccionada.
    %g maneja la grafica en la que se muestra la lección seleccionada.
    %b boton que será activado para poder reproducir la lección
    %seleccionada.
    %Función encargada de cargar la lección seleccionada en el menú
    %Cargar lecciones
    set(t,'String',leccion);
    leccion=[leccion,'.wav'];
    %Carga la lección seleccionada
    y = wavread(leccion);
    %grafica la lección seleccionada
    plot(g,y);
    axis auto;
    %Actica el botón de reproducción de la lección seleccionada
    set(b,'Enable','On');
```

#### **4.4.2 Grabación de ejercicios**

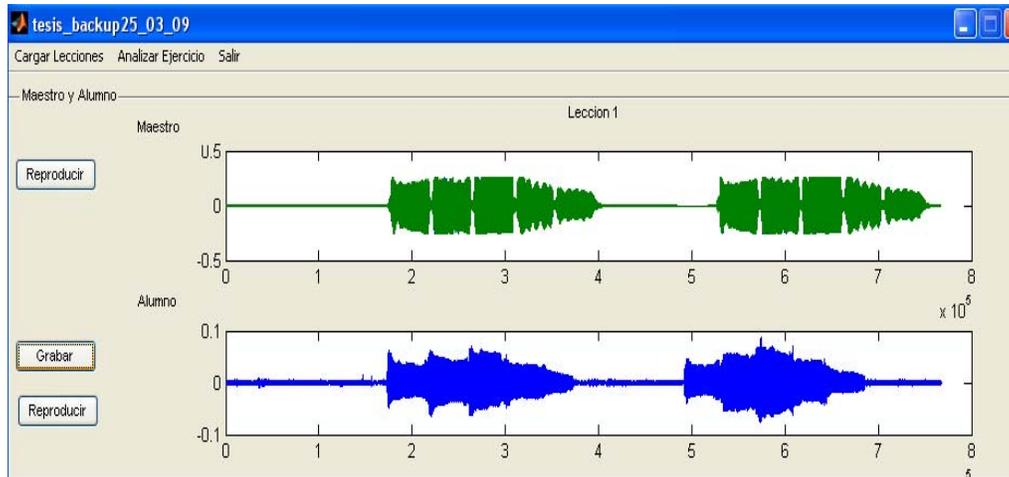
El estudiante que utilice el presente manual de canto interactivo tendrá la posibilidad de grabar su propio ejercicio basándose en la muestra pregrabada por un maestro, y contando con la ayuda de el metrónomo para mejorar la coincidencia en tiempo de cada una de los segmentos del ejercicio.

Cada ejercicio será guardado temporalmente en un archivo wav con el nombre del ejercicio para posteriormente ser comparada con lección grabada por el maestro y así realizar el análisis espectral de ambas muestras y dar como resultado la aprobación o reprobación del ejercicio.

Para que pueda ser grabado un ejercicio debe tenerse cargada una lección de maestro la cual se muestra en el marco superior de la figura 14, luego de cargarse la lección se activara el botón de grabar en el marco de trabajo del alumno. Cabe mencionar que el tiempo permitido para la grabación del ejercicio es idéntico al tiempo de grabación de la lección cargada.

En el marco inferior de la figura 14 se puede observar la gráfica tiempo versus amplitud de la lección<sup>1</sup> grabada por un alumno y a simple vista se observan errores de tiempo tanto en el inicio como en el fin de cada una de las iteraciones del ejercicio<sup>1</sup>.

Figura 14. Grabación de ejercicio 1



Las características de las lecciones grabadas por el alumno son las siguientes:

- Frecuencia de muestreo ( $F_s$ )=44100
- Número de bits por muestra=16 bits
- Número de Canales=2
- Formato=WAV

A continuación se presentan las líneas de código (Lenguaje m de Matlab) encargadas de realizar la captura de los ejercicios.

```
[eje,Fs,nBits]=wavread('Leccion 1.wav'); %Carga la grabación de la lección elegida
%Carga y reproduce el archivo de audio que contiene el metrónomo
m=wavread('metronomo.wav');
sound(m,44100)
%Graba el ejercicio con la misma duración que la Lección original
original = wavrecord(length(eje),Fs);
%graba el ejercicio a un archivo de audio WAV
wavwrite(original,Fs,nBits,'ejercicio 1');
%Grafica el ejercicio recién grabado
plot(handles.Ejercicio,original);
axis auto;
B2=handles.b2;set(B2,'Enable','On'); %activa la reproducción
```

#### 4.4.3 Aplicación de transformada de Fourier en los archivos de audio

Para obtener el espectro de la señal de audio tanto del maestro como del alumno es necesario que esta señal pase por el proceso de transformación. Esto no es más que la transformación de la señal original que se encuentra en el dominio del tiempo hacia el dominio de la frecuencia, para lograr esto se utiliza la representación en suma de senos y cosenos de la señal original haciendo uso de la FFT (Fast Fourier Transform).

Dado el tipo de aplicación que se requiere en el presente proyecto se ha creado un algoritmo que hace uso de la FFT de manera *adoc* a los requerimientos del sistema.

Para lograr la mayor exactitud en los datos obtenidos se divide la señal original en pequeños fragmentos de tamaño 4410 bits aproximadamente 1/10 segundos de grabación y se le aplica la FFT obteniendo así el vector de amplitudes luego el sistema crea el vector de frecuencias correspondiente a cada amplitud.

Para crear el vector de frecuencias se asigna a cada posición del vector el valor correspondiente dentro la siguiente serie.

1/0.1 ;2/0.1; 3/0.1; 4/0.1; ... 4410/0.1

El denominador de cada uno de los elementos utilizando en la serie anterior es el intervalo de frecuencia calculado en base a la formula

$$IF = \frac{nBits}{Fs}$$

En nuestro caso, el denominador sería 10 ya que el número de bits (nBits) es 4410 y la frecuencia de muestre (Fs) corresponde a 44100 bits/segundo.

Cuando el sistema cuenta con dos vectores (Amplitud y Frecuencia) producto de la aplicación de la FFT es necesario obtener la Frecuencia fundamental del fragmento de señal analizado.

Por definición la frecuencia fundamental se encuentra en el mayor valor obtenido en el vector de la amplitud, por tanto se deben realizar una relación punto a punto entre vectores para saber qué valor de frecuencia corresponde a la amplitud más alta en la ventana analizada.

```
Primera parte función analizar
function analizar(grafica,texto,leccion,nivel)
temp1=leccion;
[temp1,temp2]=strtok(temp1);
ejercicio=[ejercicio,temp2];
leccion=[leccion,'.wav'];
ejercicio=[ejercicio,'.wav'];
[alumno,Fsa,nBitsa]=wavread(ejercicio);
[maestro,Fs,nBits]=wavread(leccion);
i=1;
inicio=1;
n_muestrasXframe=10;
```

```

n_frames=(length(maestro)/Fs)*n_muestrasXframe;
size_frame=44100/n_muestrasXframe;
n_frames=round(n_frames);
frec_maestro=1:n_frames;
frec_alumno=1:n_frames;
while i<(n_frames)
    y_maestro=maestro((inicio):(i*size_frame));%recorta trozos de la grabacion para hacer el
    enventanado
    y_alumno=alumno((inicio):(i*size_frame));%recorta trozos de la grabacion para hacer el
    enventanado
    inicio=(i*size_frame)+1;%mueve el cursor de inicio dentro de la grabacion para ir haciendo el
    enventanado
    N=length(y_maestro);% longitud en bits de cada ventana
    T=N/Fs;    %Obtiene el intervalo de frecuencias
    k_maestro=0:N-1;
    k_alumno=0:N-1;
    freq_maestro=k_maestro/T;
    freq_alumno=k_alumno/T;
    X_maestro=fft(y_maestro)/N;%calcula la transformada y divide el vector en N para hacer
    normalizar los datos
    X_alumno=fft(y_alumno)/N; %X_maestro y X_alumno almacenan las amplitudes
    cutOff=ceil(N/2);% busca la mitad de la ventana para solo trabajar con la mitad
    X_maestro=X_maestro(1:cutOff);%recorta el vector X_maestro a la mitad
    X_alumno=X_alumno(1:cutOff);
    freq_maestro=freq_maestro(1:cutOff);
    freq_alumno=freq_alumno(1:cutOff);
    [f0_maestro,pos_maestro]=max(abs(X_maestro));%busca la mayor amplitud dentro de el
    Vector X_maestro
    [f0_alumno,pos_alumno]=max(abs(X_alumno));%busca la mayor amplitud dentro de el Vector
    X_alumno
    frec_maestro(i)=frec_maestro(pos_maestro);
    frec_alumno(i)=frec_alumno(pos_alumno);
    i=i+1;
end%fin del while

```

El anterior procedimiento se debe realizar con cada fragmento de las señales originales grabadas por el maestro y el alumno y así obtener las frecuencias fundamentales a lo largo de las grabaciones.

Posteriormente, el sistema almacenara en dos vectores diferentes las frecuencias fundamentales de la señal sonora del maestro y del alumno para ser comparadas con un algoritmo creado específicamente para esta actividad.

#### **4.4.4 Comparación de vectores de frecuencias fundamentales**

El procedimiento de comparación de los vectores de frecuencias fundamentales de las señales analizadas del maestro y del alumno es sin duda uno de los elementos más importantes de este sistema ya que permitirá obtener información acerca de la certeza con la que el alumno pudo realizar el ejercicio seleccionado y así determinar si aprobó o reprobó la práctica.

Dadas las características de las señales a comparar no es conveniente realizar una comparación casilla por casilla entre los vectores, ya que en el momento en el que se emite la voz cantada una nota puede sonar afinada mientras la frecuencia fundamental ( $F_0$ ) emitida por el alumno mantenga la condicionante siguiente.

$$F_{oAlumno} \approx \pm F_{oMaestro} * 2^{\frac{1}{2}}$$

Es decir la frecuencia fundamental del alumno puede tener una incerteza no mayor del resultado de la multiplicación de la Frecuencia fundamental emitida por el maestro multiplicada por  $2^{1/2}$  ya que al superar este limite la nota (Fo) del alumno en el ámbito musical se consideraría un semitono (bemol o sostenido) de la nota emitida por el maestro.

Otro aspecto a considerar es que en el momento de emitir una nota esta puede variar levemente en duración, momento de inicio y momento de finalización y aun así ser percibida por el oído humano como una señal semejante a una señal original dada con anterioridad.

Es por ello que para realizar la comparación de los vectores de frecuencia fundamental antes mencionados se realiza la comparación del promedio de cada 3 valores de los vectores, realizando el corrimiento de una casilla luego de cada comparación, hasta alcanzar el final de los vectores tomando en cuenta que poseen la misma longitud.

Cada una de las coincidencias encontradas en la comparación de promedios de frecuencias fundamentales es sumariada para luego verificar que el porcentaje de aciertos sea mayor que el valor estipulado para aprobar el ejercicio.

#### Segunda parte Función analizar

```
errores=0;
prom_maestro=0;
prom_alumno=0;
for i=1:length(frec_maestro)-2
    for j=i:i+2
        prom_maestro=prom_maestro+frec_maestro(j);
```

```

    prom_alumno=prom_alumno+frec_alumno(j);
end
    prom_maestro=prom_maestro/3;
    prom_alumno=prom_alumno/3;
    if(abs(prom_maestro-prom_alumno)>((prom_maestro*2^(1/12))-prom_maestro))
        errores=errores+1;
    end
end
end

```

El porcentaje de aprobación de cada ejercicio depende del nivel de análisis de ejercicio seleccionado en el menú análisis de ejercicio dentro de la aplicación, a continuación se muestra una tabla con estos datos.

**Tabla II: Niveles de dificultad en el análisis de ejercicios.**

Nivel	Porcentaje de errores
Principiantes	70%
Intermedio	60%
Avanzados	50%

```

Tercera parte función analizar
porcentaje_error=(100/(length(frec_maestro)-2))*errores;
if porcentaje_error>=50
    set(texto,'String','REPROBADO INTENTALO DE NUEVO');
    [risa,Fsa]=wavread('3.wav');
    sound(risa,Fsa);
else
    set(texto,'String','APROBADO FELICITACIONES!!!!');
    [alumno,Fsa]=wavread('0.wav');
    sound(alumno,Fsa);
end
end

```

#### 4.4.5 Mostrar e interpretar resultados

No queda duda que la interpretación de los resultados y la veracidad de los mismos es importante para poder observar avances en el proceso enseñanza aprendizaje, es por ello que la visualización e interpretación de los resultados brindados por este sistema se deben manejar de una manera adecuada.

Luego de realizar las comparaciones necesarias y el manejo de errores el sistema mostrará una grafica en la cual se verán representados en el eje “x” la posición en el vector y en el eje “y” la frecuencia fundamental encontrada.

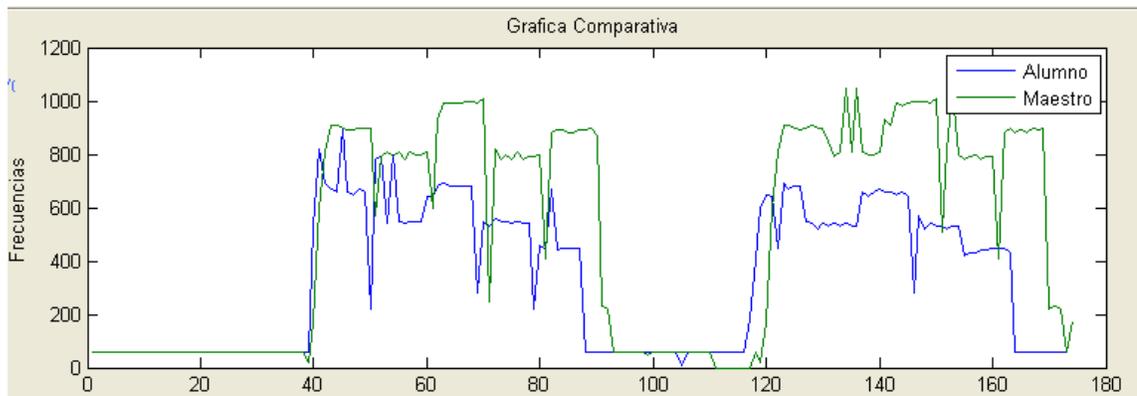
Esta visualización permitirá observar la bitácora de las frecuencias fundamentales emitidas por el maestro y el alumno y así poder observar los

Cuarta Parte función analizar

```
tiempo=1:1:length(frec_maestro);  
plot(grafica,tiempo,frec_alumno,tiempo,frec_maestro);%plotea los datos  
legend('Alumno','Maestro');%Leyendas de la grafica  
title('Grafica Comparativa ') %Titulo d la grafica  
xlabel('Posicion en el vector de Fo') %Etiqueta del eje X  
ylabel('Frecuencias Fo') %Etiqueta del eje Y
```

A continuación se presenta un ejemplo de la gráfica desplegada luego de realizado el análisis de un ejercicio de nivel principiante.

**Figura 15: Gráfica de resultados de análisis de ejercicio maestro y alumno**



En la figura 15 se muestra claramente como el alumno en la primera iteración del ejercicio inicia correctamente con el compas del maestro pero durante el ejercicio va descuadrando cada uno de las notas ejecutadas.

Otra aspecto a mencionar en esta gráfica es lo obvio que resulta el observar una grafica con picos superiores 800 mientras que la grafica inferior apenas alcanza los 700, lo cual nos permite observar que el alumno emitió notas muy por debajo de la frecuencia en la cual el maestro realizo el ejercicio.

Por último, el sistema muestra de manera clara un mensaje en el cual el alumno podrá observar si aprobó o reprobó el ejercicio para pasar a la siguiente lección o repetir la lección anterior tomando en cuenta los aspectos a mejorar observados en la gráfica.



## CONCLUSIONES

1. Toda señal análoga puede ser capturada hacia el computador y convertirse en señales digitales para luego ser analizada, sintetizada y transformada para brindar al usuario apoyo en diferentes campos de trabajo como las líneas de producción, identificación de voz, seguridad, meteorología, etc.
2. El Manual de Canto Interactivo desarrollado en el presente trabajo de fin de carrera permite practicar la entonación y el ritmo al cantar, utilizando un equipo de cómputo con una configuración mínima, esto permitirá que más personas de diferentes niveles socioeconómicos puedan interesarse en el arte del canto.
3. Este proyecto permite al alumno obtener una retroalimentación instantánea sin necesidad de contar con la presencia de un tutor o consulta de un experto y le brinda la oportunidad de analizar de manera gráfica la forma en la que son emitidos los sonidos, las diferencias entre frecuencias bajas y altas, así como la manera en la que estas se alcanzan, mejorando sustancialmente el proceso enseñanza aprendizaje del canto.
4. El algoritmo desarrollado para la comparación de las señales de audio del maestro y el alumno, permite comparar los resultados con un 85 % de exactitud debido a que las variables que intervienen en esta comparación se pueden ver afectadas por la tecnología utilizada en el momento de la grabación o sonidos ajenos a la voz del alumno en el momento de grabar el ejercicio.

5. Los tres ejercicios propuestos como introducción al canto y al manejo de la afinación y el tiempo fueron tomados de la Técnica de Canto Alemana y fueron adaptados con éxito a la aplicación creada, evaluando los ejercicios grabados por alumno con un margen de error del 15 % y teniendo una duración aproximada de 30 segundos.
  
6. Para desarrollar un algoritmo equilibrado que permitiera una aplicación que brindara resultados de manera rápida confiable en el momento de la comparación, se utilizó la *Fast Fourier Transform* como herramienta de transformación, se capturaron las señales a una frecuencia de muestreo de 441000 utilizando dos canales tomando muestras de 16 bits, además de encontrar el número exacto de muestras comparadas que no permitiera la pérdida de información.
  
7. Una señal original proveniente del dominio del tiempo puede perfectamente ser transformada al dominio de la frecuencia utilizando la transformada de Fourier y así poder encontrar la frecuencia fundamental de un segmento de señal identificando la mayor amplitud en el tiempo determinado.

## RECOMENDACIONES

1. Es importante aprobar las tres lecciones propuestas en el Manual de Canto Interactivo en los tres niveles estipulados antes de realizar ejercicios más complejos.
2. Para realizar la grabación de los ejercicios es preferible utilizar audífonos de diadema que incluyan micrófono y evitar el uso de altavoces mientras se está grabando, ya que esto podría causar el desagradable efecto feedback.
3. Antes de realizar los ejercicio de vocalización propuestos en el manual de canto es fundamental realizar la etapa de precalentamiento, para obtener mejores resultados en el momento de realizar la evaluación de los ejercicios.
4. Antes de utilizar la aplicación compruebe que los audífonos y el micrófono estén correctamente conectados y configurados además de instalar correctamente el adaptador de audio.
5. Preferiblemente no se debe abrir varias instancias de la aplicación porque esto provocará un consumo excesivo de memoria volátil, debido a que la aplicación realiza un manejo de gran cantidad de información en tiempo real.
6. Antes de realizar la grabación de un ejercicio es necesario escuchar la lección las veces que sean necesarias para memorizar el tipo de ejercicio y el tiempo en el que este es ejecutado.

7. Si se observa que la gráfica del alumno en el momento de grabar el ejercicio rellena por completo el eje "Y" del frame de gráfica grabe nuevamente el ejercicio bajando el nivel de volumen del medio de entrada seleccionado en el control de reproducción y grabación de Windows.

## BIBLIOGRAFÍA

1. De León, Moisés. **Documento respiración y vocalización.**  
México, Arquidiócesis Primada de México, Pastoral Grupos Musicales Parroquiales Juveniles, s.a. 4 pp.
2. Garcia, Daniel. **Método Aleman de Canto.** Argentina, Mystical-Sword. 2001. 124.pp.
3. López, Ernesto y otros. **Búsqueda de música por tarareo.**  
Proyecto de fin de carrera, Uruguay, Universidad de la República Oriental de Uruguay, Facultad de Ingeniería, Instituto de Ingeniería Eléctrica, 2004. 97.pp.
4. Mosquera River, Ismael. **Sistema de conversión de voz a Midi.**  
Proyecto Fin de Carrera, Universidad Pompeu Fabra, España, Ingeniería Técnica en Informática de Sistemas 2004. 98 pp.
5. Ortiz, Luís **Identificación automática de acordes musicales.**  
Tesis doctoral, Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros de Telecomunicaciones, 2002. 294.pp.
6. Quintero, Antonio. **Tutorial básico para el manejo de señales con Matlab.** Argentina, Universidad Nacional de Quilmes, s.a.  
<http://musica.unq.edu.ar/personales/ebonnier/cam2/matab/tutorial2632.html>, visitado el 20 de enero de 2009.
7. Stein, Hanns. **El arte de cantar: su dimensión cultural y pedagógica.**  
Chile, Facultad de Artes, Universidad de Chile, Revista musical Chilena, 54 (194): 2000.
8. The MathWork.Inc. **Ayuda del Producto Matlab 7**  
<http://www.mathworks.com>, visitado el 4 de febrero de 2009.
9. The MathWork.Inc **MatLab 7 Getting Started Guide.**  
<http://www.mathworks.com>, visitado el 5 de febrero de 2009.



## **ANEXOS**

### **Manual de instalación**

### **Requerimientos mínimos**

#### **Hardware**

- Procesador Intel Pentium, Intel Celeron, AMD.
- 512 Mb memoria Ram.
- 100 Mb. de espacio físico en disco duro (15 Mb para configuración e instalación).
- Tarjeta de audio.
- Adaptador y monitor de vídeo Super VGA (800 × 600) o de mayor resolución
- Kit Audífonos y micrófono.
- Unidad de CD-ROM o DVD.
- Ratón y teclado.

#### **Software**

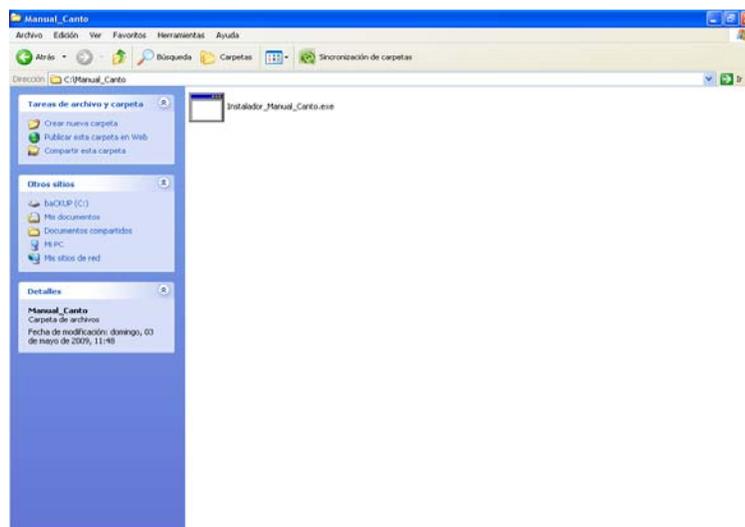
- Microsoft Windows Xp, Vista.
- Adobe Reader Versión 7 o superior.
- Drivers de tarjeta de audio incorporado o insertada.

## Instalación

A continuación se detallan los pasos para la instalación del Manual de Canto Interactivo:

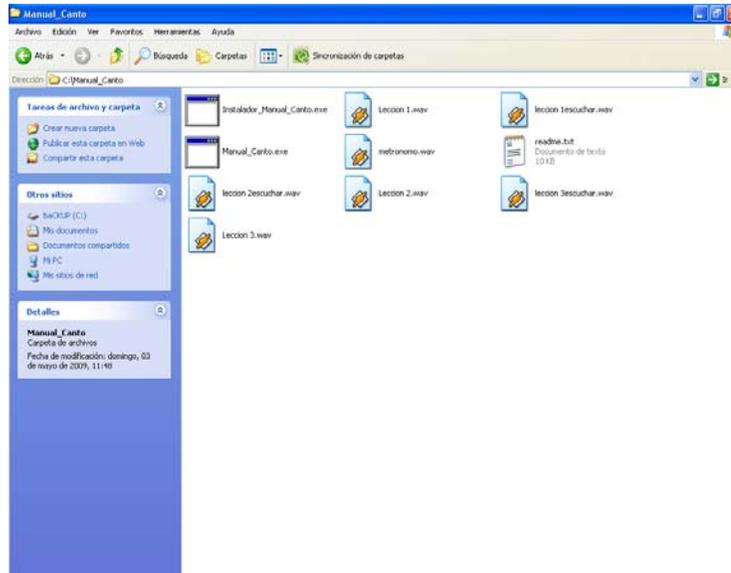
1. Abrir la bandeja del CD ROM del computador.
2. Introducir el disco del Manual de Canto Interactivo.
3. Copiar el contenido la carpeta Manual\_Canto del CD en la ubicación en la que desea instalar el programa.
4. Localizar el archivo Instalador\_Manual\_Canto.exe dentro de la carpeta Manual\_Canto y ejecutarlo.

**Figura 16. Ubicación del archivo instalador Manual\_Canto.exe**



5. El programa desempaquetara y configurara los archivos necesarios para la ejecución del sistema (Archivos exe, wav, txt, pdf) de Manual de Canto Interactivo.

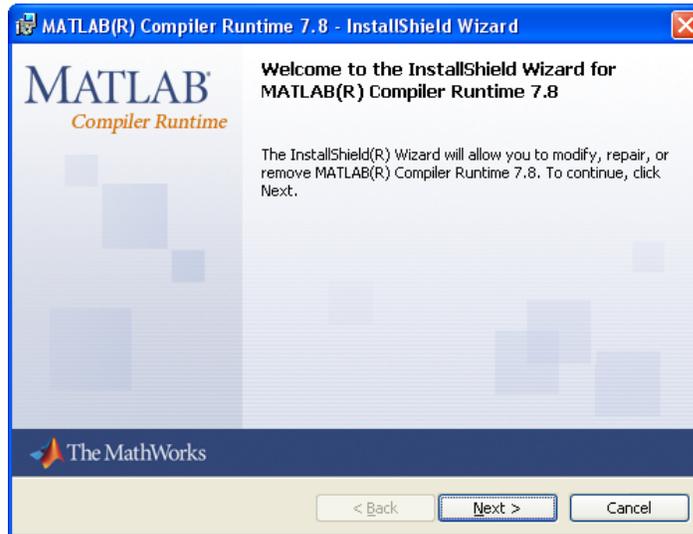
Figura 17. Archivos utilizados para la ejecución de la aplicación.



6. Antes de ejecutar el archivo `.manual_de_canto_interactivo.exe` es necesario instalar el archivo Compiler Runtime de Matlab `MCRInstaller.exe` para que la aplicación pueda correr en la máquina en la que se va a instalar.

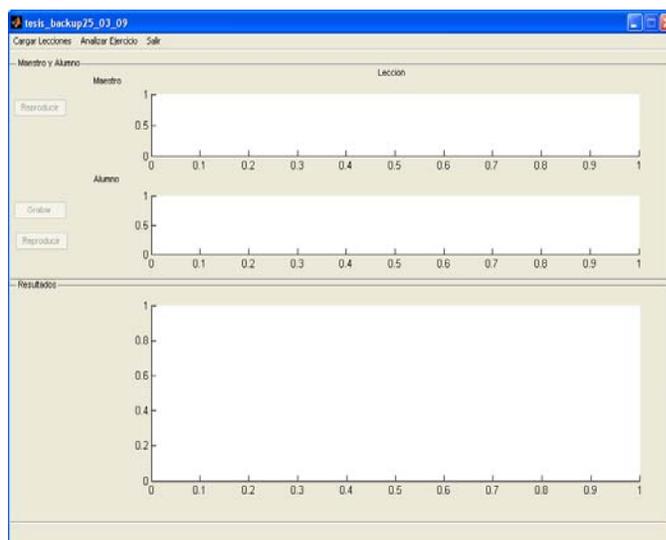
**NOTA:** Automáticamente se detectará si se debe instalar el archivo `VCREDIST_X86` antes de instalar `MCRInstaller.exe`.

Figura 18. Pantalla inicial de instalación de MCRInstaller.exe



7. Buscar y ejecutar el archivo Manual\_Canto\_Interactivo.exe el cual permitirá abrir la aplicación del Manual de Canto Interactivo.

Figura 19. Aplicación "Manual de Canto Interactivo".



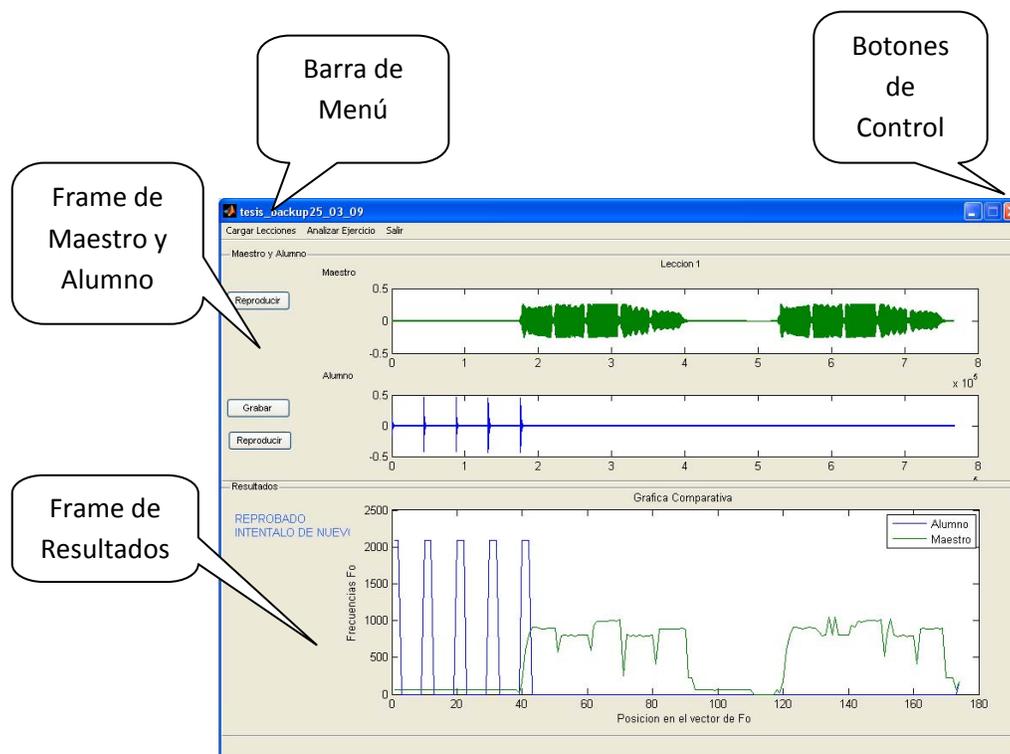
## Manual de Usuario

### Entorno de trabajo

Al ejecutar el archivo Manual\_Canto.exe este dará acceso a la ventana de trabajo del manual de canto interactivo. Esta ventana está dividida en dos frames principales que llevan como título “Maestro y alumno” y “Resultados”.

En la esquina superior izquierda de la venta se encuentra ubicada la barra de menú principal de la aplicación y en la esquina superior derecha se encuentra el menú de botones de control.

Figura 20. Entorno de trabajo de “Manual de Canto Interactivo”.



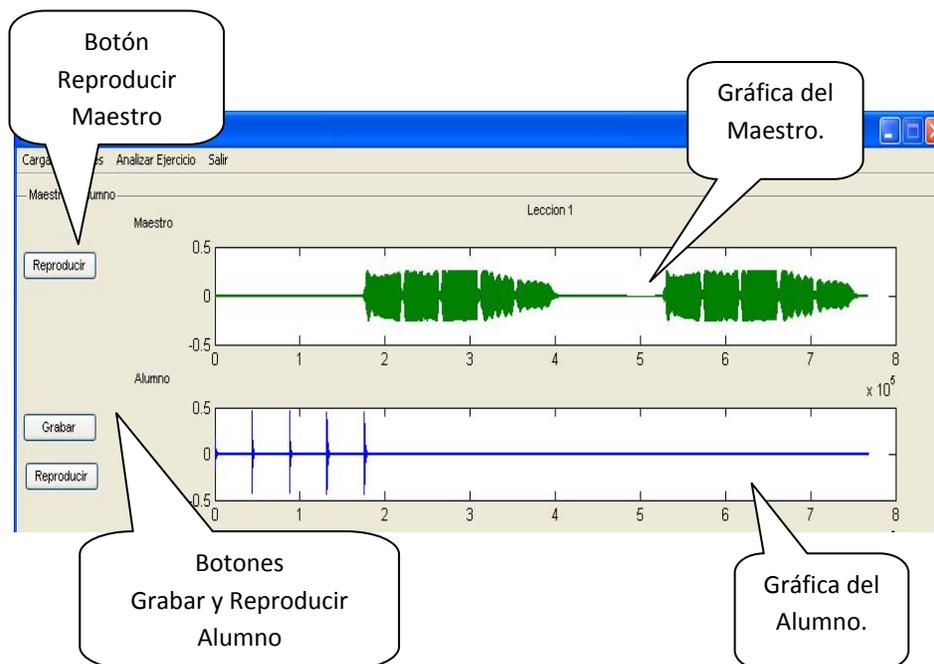
## **Frame maestro y alumno**

Este frame permite visualizar la gráfica tiempo versus amplitud de la lección seleccionada (maestro) y del ejercicio grabado por el alumno.

Al lado izquierdo de la grafica del maestro se encuentra el botón reproducir que permite escuchar la lección seleccionada grabada por el maestro.

Al lado izquierdo de la gráfica del alumno se encuentra el botón grabar que permite al alumno grabar la lección luego de escuchar la muestra del maestro. Además debajo de este botón se encuentra el botón reproducir que permite reproducir la lección grabada por el alumno, este botón se activa en el momento en el que el alumno termina de grabar el ejercicio.

**Figura 21. Frame maestro y alumno**



## Menú principal

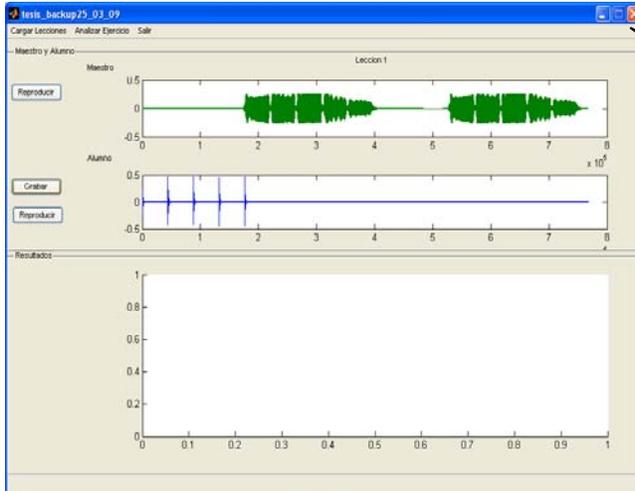
El menú principal presenta las siguientes tres opciones:

- Cargar lección: Permite seleccionar cualquiera de las tres lecciones cargadas a la presente versión del manual de canto interactivo.
- Analizar ejercicio: Permite seleccionar el grado de dificultad con el cual será evaluado el ejercicio, presentando los niveles principiante, intermedio y avanzado.
- Salir: Cierra el manual de canto interactivo.

## Botones de control

Los botones de control se encuentran ubicados en la esquina superior derecha ver figura 22.

**Figura 22. Ubicación botones de control.**



Botones de Control:  
-Maximizar.  
-Minimizar.  
-Cerrar.

- Botón minimizar: Minimiza la venta del manual de canto y la convierte en un icono en la barra de Tareas permaneciendo activa la aplicación.
- Botón maximizar: La ventana pasa a ocupar toda la pantalla
- Botón restaurar: Éste aparece cuando la ventana está maximizada y sirve para devolver a ésta su tamaño anterior, haciendo clic en este botón.
- Botón cerrar: Cierra totalmente la ventana, que equivale a Salir de la aplicación (programa) liberando los recursos del sistema que utilizaba,
- Mover una ventana: La forma más rápida es colocar el puntero del mouse sobre la barra de título de la ventana, haciendo clic con el botón izquierdo del mouse, sin soltarlo y arrastrarlo hasta la posición deseada y libere el botón.<sup>56</sup>

## Uso de la aplicación

### Cargar lección

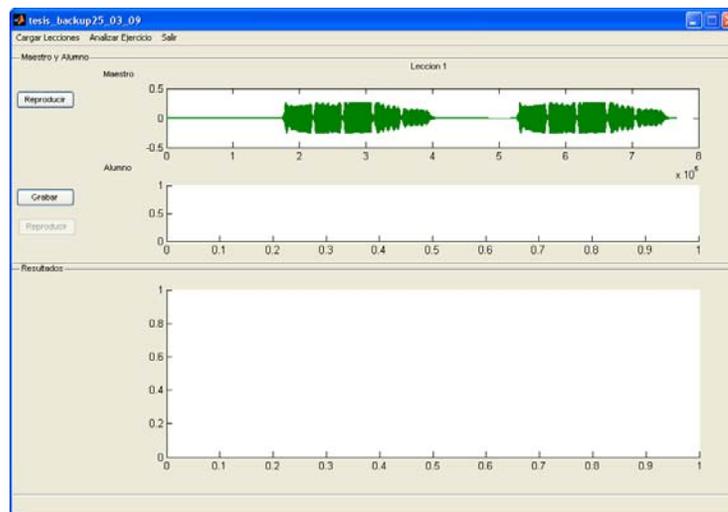
<sup>56</sup> Guía para Windows 98. Disponible en: <http://anderground.galeon.com/cvitae405772.html>

Al cargar la aplicación todas las gráficas aparecerán vacías y se utilice la aplicación estas mostrarán la información requerida por el usuario.

El primer paso para utilizar el manual de canto interactivo es seleccionar la lección a realizar, en el primer submenú “Cargar Lecciones” se puede escoger una de tres lecciones.

Luego de seleccionar la lección se cargara en el área de gráfica de maestro la gráfica en color verde correspondiente a la lección seleccionada, inmediatamente también se activara el botón reproducir para que el alumno pueda escuchar la muestra grabada por el maestro como se observa en la figura 23.

**Figura 23. Gráfica de la lección 1.**



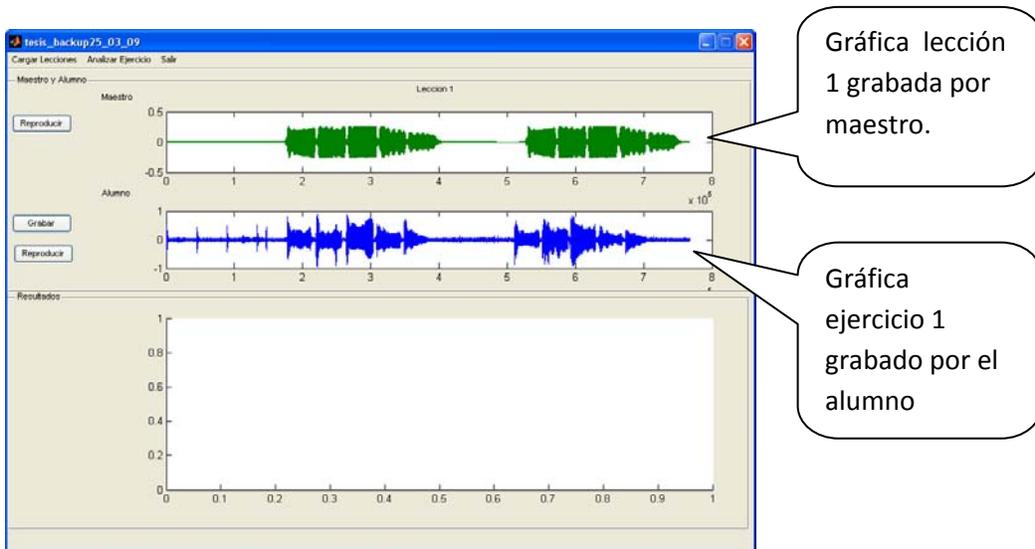
**Grabar ejercicio**

Luego de cargar la lección el botón grabar del área de alumno se activará para que se inicie la grabación del ejercicio.

Al pulsar el clic izquierdo del mouse sobre el botón grabar se escucharan el sonido del metrónomo que ayudará al alumno a llevar el compas del ejercicio, el ejercicio se debe ejecutar luego de contar los primeros cuatro tiempos del metrónomo.

La duración de grabación del ejercicio durará exactamente el tiempo de duración que posea la lección seleccionada. Al finalizar la grabación se mostrará en color azul la gráfica creada a partir de la señal de entrada grabada por el alumno, tal como se muestra en la figura 23.

**Figura 24. Gráfica de grabación de ejercicio realizado por el alumno.**



Luego de grabado el ejercicio se activara el botón reproducir para que el alumno pueda escuchar el ejercicio grabado antes de proceder a ejecutar el análisis y así repetir la grabación del ejercicio si el alumno lo considerara necesario.

### **Analizar ejercicio**

El menú analizar ejercicio permite seleccionar el grado de dificultad con el cual será evaluado el ejercicio, existen tres opciones: principiante, intermedio y avanzado.

Al dar clic sobre el nivel de análisis se iniciara el proceso de análisis que durará algunos segundos dependiendo de las características del equipo con que se cuente.

### **Interpretar resultados**

En el frame de resultados se podrá observar un área de gráficas en la cual se visualizarán las gráficas de frecuencias fundamentales de las señales de maestro (color verde) y alumno (color azul).

En esta gráfica es fácil identificar los momentos en los cuales se encuentra un desfase en tiempo y en afinación estos resultados gráficos deberían de coincidir en la mayoría de los puntos a lo largo del eje x para poder darse como aprobado el ejercicio.

La gráfica está acompañada de un mensaje en el cual se da por aprobado o reprobado el ejercicio.

Para poder repetir el ejercicio no es necesario cargar nuevamente la aplicación ni la lección únicamente se debe grabar de nuevo el ejercicio y realizar el análisis deseado, para volver a revisar los resultados.