

**Universidad Central “Marta Abreu” de Las Villas**

**Facultad de Ingeniería Eléctrica**

**Centro de Estudios de Electrónica y Tecnologías de la Información**



## **TRABAJO DE DIPLOMA**

# **Interfaz gráfica para manipular y procesar archivos de audio de voces patológicas.**

**Autor: Víctor Manuel Del Sol Fariñas**

**Tutor: MSc. Hector Arturo Kairuz Hernández-Díaz**

**Santa Clara**

**"Año 57 de la Revolución"**

**"2015"**

**Universidad Central “Marta Abreu” de Las Villas**

**Facultad de Ingeniería Eléctrica**

**Departamento de Automática y Sistemas Computacionales**



## **TRABAJO DE DIPLOMA**

# **Interfaz gráfica para manipular y procesar archivos de audio de voces patológicas.**

**Autor: Víctor Manuel Del Sol Fariñas**

([vdelsol@uclv.edu.cu](mailto:vdelsol@uclv.edu.cu))

**Tutor: MSc. Hector Arturo Kairuz Hernández-Díaz**

([akairuz@uclv.edu.cu](mailto:akairuz@uclv.edu.cu))

**Santa Clara**

**"Año 57 de la Revolución"**

**"2015"**



Hago constar que el presente trabajo de diploma fue realizado en la Universidad Central “Marta Abreu” de Las Villas como parte de la culminación de estudios de la especialidad de Ingeniería en Biomédica, autorizando a que el mismo sea utilizado por la Institución, para los fines que estime conveniente, tanto de forma parcial como total y que además no podrá ser presentado en eventos, ni publicados sin autorización de la Universidad.

---

Firma del Autor

Los abajo firmantes certificamos que el presente trabajo ha sido realizado según acuerdo de la dirección de nuestro centro y el mismo cumple con los requisitos que debe tener un trabajo de esta envergadura referido a la temática señalada.

---

Firma del Tutor

---

Firma del Jefe de Departamento  
donde se defiende el trabajo

---

Firma del Responsable de  
Información Científico-Técnica

## PENSAMIENTO

*"No se le puede enseñar nada a un hombre; sólo se le puede ayudar a encontrar la respuesta dentro de sí mismo".*

*Galileo Galilei*

**DEDICATORIA**

*A mis adorados y excelentes padres María E. y Andrés*

*A mí querida familia*

*A mi novia Dailén*

## AGRADECIMIENTOS

*A mis compañeros de estudio y a mis profesores del Departamento de Biomédica, quienes directa o indirectamente han contribuido en mi trabajo investigativo.*

*A mis abuelos, tías, tíos, mis amigas y amigos con quienes he podido contar bien en el aspecto profesional o personal.*

*A todas aquellas personas que me han apoyado en este empeño, con un gesto, una mirada, una interrogante, un pensamiento positivo; mi especial y nunca olvidado agradecimiento.*

### **AGRADECIMIENTOS ESPECIALES**

*A mi tutor MSc. Hector Arturo Kairuz Hernández-Díaz por su incondicional apoyo, asesoramiento, dedicación y consagración, quien ha sido imprescindible en este proceso investigativo.*

*A mi novia Dailén que a pesar de encontrarnos en iguales procesos, ha constituido un sostén en el que me ha apoyado para poder continuar y culminar esta meta que aunque alcanzable es pedregoso su camino.*

*A mis padres Andrés Del Sol Garay y María E. Fariñas Herrera; los dos, inspiración de mi vida, ejemplo de laboriosidad y profesionalidad. A los dos, mi infinito amor y agradecimiento por su sacrificio, comprensión, apoyo macro, amor, cariño y entrega en todo momento.*

## TAREA TÉCNICA

1. Análisis crítico del estado del arte sobre el análisis acústico de voces patológicas.
2. Creación de un modelo entidad relación óptimo para una base de datos de pacientes disártricos.
3. Diseño de una interfaz gráfica para el manejo de la base de datos.
4. Evaluación de los resultados obtenidos por los especialistas en la práctica clínica.
5. Elaboración de recomendaciones sobre el uso de la interfaz gráfica.

---

Firma del Autor

---

Firma del Tutor



## **RESUMEN**

La voz contiene información útil para el diagnóstico médico, además de constituir una rama de la medicina (Logopedia y Foniatría). En numerosos países se han creado bases de datos que recopilan la información de los pacientes, las valoraciones de los especialistas, las grabaciones realizadas además de otras medidas. Todo esto para el análisis y tratamiento de disímiles patologías de habla y voz. Por otra parte también se utiliza en la práctica clínica para llevar registro de los pacientes y como realimentación del tratamiento. En nuestra localidad no se cuenta con una aplicación apropiada que permita reunir la información antes mencionada en una base de datos. La presente investigación se propone diseñar e implementar una base de datos y un software para su manejo. Este debe cumplir los requerimientos básicos propuestos por los usuarios (Especialistas de la Consulta de Logopedia y Foniatría). La aplicación finalmente debe ser instalada y evaluada en la consulta provincial del Servicio de Logopedia y Foniatría de la provincia de Villa Clara.

## TABLA DE CONTENIDOS

PENSAMIENTO .....	i
DEDICATORIA .....	ii
AGRADECIMIENTOS .....	iii
TAREA TÉCNICA .....	v
RESUMEN .....	vi
INTRODUCCIÓN .....	1
1. Fundamentos Teóricos .....	4
1.1. Características de la señal de voz .....	4
1.1.1 Señal sonora .....	4
1.1.2 Señal no sonora .....	5
1.1.3 Plosivas .....	5
1.2 Modelo de producción de la voz .....	6
1.2.1 Causas de los trastornos del lenguaje oral. ....	8
1.2.2 Un factor causal de alteraciones del lenguaje es el cronológico.....	8
1.2.3 Análisis Acústico en Voces patológicas .....	9
1.3 Base de datos.....	11
1.3.1. Base de datos fonética, Modelo 4332. ....	12
1.3.1.1 Manual .....	13

1.3.2 Base de datos de voces patológicas y Programa Modelo 4337., .....	13
1.3.2.1 La base de datos .....	14
1.3.2.2 Características de la base de datos de voces patológicas, modelo 4337.....	14
1.3.3 Aplicaciones.....	15
1.4. Software para el análisis de las señales de voz dentro de bases de datos. ....	15
1.4.1. Praat .....	15
1.4.2. EMU.....	16
1.4.2.1 Estructura actual del EMU.....	17
Capítulo 2. Materiales y Métodos:.....	18
2.1. Diseño de la Base de Datos: .....	18
2.1.1. Situación problemática: .....	18
2.2. Diseño del Software.....	20
2.3. El lenguaje SQL.....	23
2.4. Net Framework .....	24
2.4.1 Estructura interna del entorno de ejecución en lenguaje común. ....	25
2.4.2 Características .....	26
2.5. Microsoft Visual Studio.....	26
2.5.1. Versiones .....	27
2.5.1.1 Visual Studio 6.0.....	27
2.5.1.2 Visual Studio .NET (2002) .....	28
2.5.1.3 Visual Studio .NET 2003.....	28
2.5.1.4 Visual Studio 2005.....	29
2.5.1.5 Visual Studio 2008.....	29
2.5.1.6 Visual Studio 2010.....	30

2.6. C Sharp (C#) .....	30
2.6.1. Origen .....	30
2.6.2. En la actualidad existen los siguientes compiladores para el lenguaje C#. ....	31
2.6.3. El estándar ECMA-334 lista las siguientes metas en el diseño para C#: .....	32
CAPÍTULO 3: RESULTADOS Y DISCUSIÓN. ....	33
3.1. Software Pafcont. ....	33
3.1.1.Instalación de la Aplicación:.....	33
3.1.2. Modo de empleo de la aplicación .....	34
3.1.3. Manejo de la Aplicación. ....	35
3.1.3.1. Introducir un nuevo paciente a la base de datos. ....	35
3.1.3.2. Ingresar un diagnóstico a un paciente previamente seleccionado. ....	36
3.1.3.3. Modificar un paciente. ....	37
3.1.3.4. Realizar Búsquedas.....	38
3.1.3.5 Operaciones con Archivos de Audio. ....	39
Conclusiones Parciales .....	41
CONCLUSIONES Y RECOMENDACIONES .....	42
Conclusiones:.....	42
Recomendaciones .....	43
REFERENCIAS BIBLIOGRÁFICAS .....	44

## **INTRODUCCIÓN**

El habla es la forma natural de comunicación del hombre y constituye la principal diferencia entre la especie humana y el resto de los animales. El lenguaje, se ha convertido en un campo de gran cantidad de investigaciones para la medicina y las aplicaciones de la tecnología. Los conocimientos desarrollados acerca de la fisiología de la voz y su procesamiento al campo de las ciencias de la vida, tienen gran utilidad en sistemas de diagnóstico y constituyen la base de numerosas terapias de rehabilitación, sin embargo, por constituir la voz una señal compleja de la cual pueden extraerse muchos parámetros con el objetivo de modelarla o caracterizarla, este tema ha sido tan explotado como desconocido en plenitud [1].

Por ello el desarrollo alcanzado en la medicina moderna ha hecho que el personal médico necesite cada vez más de algoritmos y equipos especializados para el diagnóstico y tratamiento de diferentes enfermedades. Las técnicas automáticas que permiten el diagnóstico de enfermedades de la voz están siendo cada vez más utilizadas por el personal de salud y el producto de su utilización cuenta con el aval de la comunidad científica moderna. El uso de técnicas de procesamiento digital de señales en la caracterización de señales de voz permite obtener medidas objetivas de los diferentes rasgos que caracterizan a estas señales biológicas y reduce la subjetividad introducida por la apreciación humana [2, 3].

En las consultas de Logopedia y Foniatría las evaluaciones subjetivas han sido la herramienta principal para la clasificación de las distintas patologías [4], por lo que surgen dudas sobre la confiabilidad y la validez de las mismas, especialmente por

el hecho de que el proceso terapéutico y su seguimiento se lleven a cabo por técnicos y personal médico con escasos recursos especializados, otro de los problemas pero esta vez en la consulta de Villa Clara es la carencia de una herramienta que permita la organización de la información, incluido los archivos de audio, de los pacientes con diferentes patologías que a diario asisten a esta consulta de ahí que la tarea planteada en este trabajo se enmarca en las investigaciones que se llevaron a cabo en el Laboratorio de Procesamiento de Voz, perteneciente al Centro de Estudios de Electrónica y Tecnologías de la Información (CEETI) con la finalidad de encontrar soluciones a esta traba y así garantizar un trabajo más eficiente y organizado, de ahí que el problema científico de nuestro trabajo sea el siguiente:

### **Problema científico**

En la consulta provincial de Logopedia y Foniatría de Villa Clara existen diferentes estudios investigativos relacionados con las patologías de habla y voz. Pero este importante servicio médico carece de una base de datos que reúna los pacientes, los diagnósticos, observaciones y los archivos de audio. Esta situación es una limitante para realizar investigaciones o análisis, además puede provocar la pérdida de información.

Para solucionar esta problemática se presenta el siguiente objetivo general.

### **Objetivo general**

Desarrollar una aplicación que permita generar y modificar una base de datos de pacientes con el fin de utilizar la misma con propósitos clínicos e investigativos.

Para el cumplimiento del objetivo general se han planteado los siguientes objetivos específicos.

**Objetivos Específicos**

Estudiar con los especialistas médicos las necesidades y requerimientos que permitan el diseño de la base de datos.

Desarrollar una aplicación lo más amena posible a los usuarios finales.

Organización del informe

El informe de la investigación se estructura en introducción, capitulario, conclusiones y recomendaciones, referencias bibliográficas.

En la introducción se deja definida la importancia, actualidad y necesidad del tema que se aborda.

En el Capítulo 1 se presentan los principales conceptos que se tratan en este trabajo así como las características de la señal de voz, además se brinda una panorámica general existente en torno al problema que se aborda y que motiva este trabajo.

En el Capítulo 2 se explica el diseño metodológico de la investigación y el modelo desarrollado para la confección de esta herramienta (software) ya con las modificaciones efectuadas para utilizarlo en voces patológicas.

En el Capítulo 3 se presentan los resultados de esta herramienta.

## **1. Fundamentos Teóricos**

### **1.1. Características de la señal de voz**

En este capítulo abordaremos todo lo relacionado con las características de la señal de voz y programas para el análisis de bases de datos.

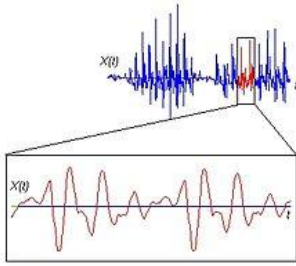
La voz humana se produce voluntariamente por medio del aparato fonatorio. El mismo está formado por los pulmones como fuente de energía en la forma de un flujo de aire, la laringe, que contiene las cuerdas vocales, la faringe, las cavidades oral (o bucal) y nasal y una serie de elementos articulatorios: los labios, los dientes, el alvéolo, el paladar, el velo del paladar y la lengua [5].

Básicamente, la Señal de Voz puede clasificarse en los siguientes tipos, Sonora, No Sonora y Plosiva [6].

#### **1.1.1 Señal sonora**

La señal sonora, Figura 1 se genera por la vibración de las cuerdas vocales manteniendo la glotis abierta, lo que permite que el aire fluya a través de ella. Estas señales se caracterizan por tener alta energía y un contenido frecuencial en el rango de los 300 Hz a 4000 Hz presentando cierta periodicidad como se observa en la, es decir son de naturaleza cuasi periódica. El tracto vocal actúa como una cavidad resonante reforzando la energía en torno a determinadas frecuencias (formantes). Toda las vocales se caracterizan por ser sonoras pero existen consonantes que también lo son, tales como, la /b/, /d/ y la /m/, entre otras [6].

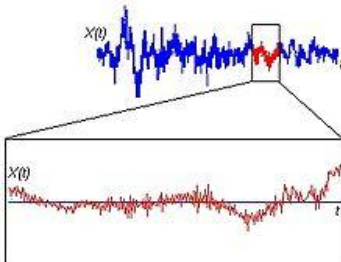




**Figura 1. Señal sonora**

### 1.1.2 Señal no sonora

A esta señal Figura 2 también se le conoce como señal fricativa o sorda, y se caracteriza por tener un comportamiento aleatorio en forma de ruido blanco. Tienen una alta densidad de cruces por cero y baja energía comparadas con las señales de tipo sonora. Durante su producción no se genera vibración de las cuerdas vocales, ya que, el aire atraviesa un estrechamiento, y genera una turbulencia. Las consonantes que producen este tipo de sonidos son la /s/, la /f/ y la /z/ entre otras[6] .



**Figura 2. Señal no sonora**

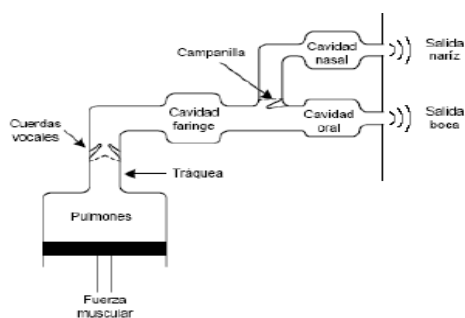
### 1.1.3 Plosivas

Esta señal se genera cuando el tracto vocal se cierra en algún punto, lo que causa que el aire se acumule para después salir expulsado repentinamente (explosión). Se caracterizan por que la expulsión de aire está precedida de un silencio. Estos sonidos se generan por ejemplo, cuando se pronuncia la palabra '*campo*'. La /p/ es una consonante de carácter plosivo, y existe un silencio entre las sílabas '*cam*' y

'po'. Otras consonantes que presentan esta característica son /t/, y /k/, entre otras[6].

## 1.2 Modelo de producción de la voz

La señal de voz se comienza a generar en el cerebro el cual elabora el mensaje, posteriormente este envía impulsos nerviosos a cada órgano motor relacionado en este proceso, los pulmones son los protagonistas de este, los cuales almacenan el aire, cuyo aire es comprimido por el diafragma y debido a esta presión es producido el sonido el cual después se desplaza desde alguna abertura en la cuerdas vocales hasta los labios. A continuación se observa todo el proceso en la Figura 3.



**Figura 3. Modelo funcional de generación de voz.**

Se define como patología aquella enfermedad relacionada con las anomalías del organismo, lo "patológico. La patología del lenguaje, se refiere a los trastornos o desórdenes del lenguaje para ello el especialista en lenguaje necesita por ende, desarrollar un sentido de observación y análisis de los síntomas, para lograr encargarse de la rehabilitación funcional de dichos trastornos y al médico le corresponde el tratamiento orgánico de los mismos; cada profesional es insustituible en su campo para la corrección de los trastornos del lenguaje [7].

La habilidad para aprender el lenguaje oral es un rasgo esencial del ser humano, cuya adquisición se basa en una variedad de mecanismos biológicos dados, así como en experiencias concretas y personales. Por ello, uno de los tópicos más importantes en la investigación y detección de las dificultades lingüísticas es el

estudio de las alteraciones del lenguaje que pueden comprometer, tanto la comprensión como la producción oral.

El concepto de lenguaje “normal”, reúne una serie de características, ajustadas a las normas sociales, que no obstaculizan las relaciones entre los individuos que forman la colectividad y no coexiste una imposibilidad verdadera de expresión. Todos los rasgos que se oponen o salen de este concepto entran en el campo de la patología. Por tanto, las anomalías del lenguaje son, todas las diferencias de la norma en cuanto a forma, grado, cantidad, calidad, tiempo y ritmo lingüístico que dificultan las posibilidades de expresión interpersonal y que implican una deficiencia duradera de la habilidad lingüística.

El concepto de trastorno de lenguaje oral es muy amplio e impreciso. Terminológicamente ha sido expresado de muy diversas maneras: dificultad, desorden, alteración o perturbación lingüística.

También, puede definírsele, como un conjunto de elementos semiológicos que suponen una comparación de las estructuras lingüísticas de un sujeto con las de otros sujeto que habla normalmente, utilizando como referencia la posesión y buen funcionamiento de los elementos anatómicos (normalidad de los órganos lingüísticos), de los elementos funcionales (contexto socializador y lingüístico) y de los elementos ambientales (entorno comunicativo de continuo estímulo).

La Asociación Americana de Habla — Lenguaje — Audición (ASHA, por sus siglas en Inglés), define los trastornos del lenguaje como un trastorno en la adquisición, comprensión o expresión normal del lenguaje hablado o escrito. El trastorno puede implicar a todos o algunos de los componentes: fonológico, morfológico, semántico, sintáctico o pragmático del sistema lingüístico [7].

Un trastorno específico del lenguaje, es aquella dificultad del funcionamiento lingüístico que se encuentra en personas sin afectaciones neurológicas o sensoriales detectables, con una inteligencia no verbal dentro de los límites de la normalidad y que, a pesar de una estimulación adecuada, una educación suficiente y unos progresos observables, no llegan a alcanzar un nivel de

conocimiento lingüístico que les permita comprender y expresarse de forma igualmente correcta en toda la gama posible de situaciones comunicativas.

Resumiendo, puede anotarse entonces que toda alteración del lenguaje oral puede comprometer tanto la comprensión como la expresión del mismo, por consiguiente, afectará la formación de conceptos, pensamiento, aprendizaje formal, la interacción social y por consiguiente la autoestima.

### **1.2.1 Causas de los trastornos del lenguaje oral.**

Las causas que generan los trastornos lingüísticos orales, se abarcan desde diferentes perspectivas o enfoques [7]:

En primer término están la etiológicas (en función de su origen / la etiología es el estudio de las causas que originan los padecimientos) que se refieren a causas de orden genético, neurológico o anatómico, como las malformaciones de los órganos del habla, las deficiencias auditivas, motoras o de orden neurológico.

En segundo lugar, están las causas funcionales (dificultades tradicionalmente consideradas de tipo psicológico) que no afectan a la comunicación social y afectiva, pero comprometen el aprendizaje lingüístico. Son generalmente defectos en el proceso fisiológico de los sistemas que intervienen en la emisión de la palabra, aunque los órganos se encuentren en perfecto estado.

### **1.2.2 Un factor causal de alteraciones del lenguaje es el cronológico.**

En función de este factor se establece una dicotomía entre alteraciones adquiridas y alteraciones en la adquisición del lenguaje oral. Las adquiridas pueden tener una causa conocida de origen orgánico y sobrevienen cuando el lenguaje está suficientemente adquirido como ocurre, por ejemplo, con los trastornos afásicos [7].

En relación con la dicotomía lenguaje – habla, la afectación de la comprensión y la expresión llegarían a plantear una alteración del lenguaje y la afectación de la producción dirigiría a una problemática específica del habla. Las dificultades del lenguaje son más graves, en principio, que las del habla y se refieren a los

aspectos más externos de la comunicación, como cuestiones de voz, ritmo y articulación.

Algunos investigadores atribuyen una base de transmisión genética a las dificultades lingüísticas, mientras que otros señalan como causa de éstas la influencia de factores socioculturales. Sin embargo, no se descarta una combinación de ambas causales, puesto que generalmente las primeras inciden sobre las segundas y viceversa [7].

### **1.2.3 Análisis Acústico en Voces patológicas**

La presencia de patologías en las cuerdas vocales causa cambios significativos en los patrones vibratorios normales de las mismas, lo cual impacta en la calidad resultante de la producción de voz.

Existen diversos procedimientos de rutina para examinar la laringe con propósitos clínicos o de investigación, los cuales incluyen laringoscopia fibroscópica rígida y flexible (entubación del paciente con un instrumento de fibra óptica), videoestroboscopia (iluminación estroboscópica de la laringe, útil para la visualización de movimientos), electromiografía (observación indirecta del estado funcional de la laringe) y videofluoroscopia (técnica radiográfica en la cual se hacen ingerir al paciente distintos bolos de bario, como sustancia de contraste). Sin embargo, existen también limitaciones para llevar a cabo estos estudios, como ser: 1) la laringe está situada fuera de la vista, en un lugar profundo del cuello; 2) el interior de la laringe es oscuro y debe ser adecuadamente iluminado para el examen; 3) los movimientos de las cuerdas vocales durante la fonación son demasiados rápidos para ser capturados por un sistema óptico convencional [8].

El desarrollo de las nuevas tecnologías de la información ha impactado directamente en la mayoría de los aspectos de la vida moderna. En la esfera de la salud, el impacto ha sido de gran relevancia dando lugar al desarrollo de innumerables herramientas y equipos para el diagnóstico, tratamiento y rehabilitación. El estudio de la voz mediante el análisis acústico ha sido objeto de innumerables investigaciones, destacándose el nacimiento de varias herramientas

para el estudio de diferentes aspectos del proceso de producción de la voz que no son fácilmente percibidos de forma perceptual o con la ayuda de dispositivos mecánicos sencillos. Una de las herramientas más utilizadas y comercializadas para el análisis de voces patológicas es el CSpeech, desarrollado por KayElemetrics [9]. Dicha herramienta es de propósito general y permite obtener varios parámetros de la voz que reflejan muchas de las perturbaciones escuchadas [10].

En los últimos años ha crecido el interés por el análisis acústico de voces normales y patológicas como método alternativo de diagnóstico. En este tipo de estudio se aplican diferentes técnicas de procesamiento de señal: procesamiento temporal y frecuencial para la obtención de parámetros acústicos [9], y extracción de la señal glótica de presión sonora mediante modelos de la dinámica del aire en la laringe a partir de filtros inversos, análisis de predicción lineal, etc. [11, 12]. El análisis acústico demuestra ventajas sobre los estudios anteriormente citados debido a su naturaleza no invasiva y a su potencial para proveer datos cuantitativos acerca del estado clínico de las funciones de la laringe y del tracto vocal, con adecuados tiempos de análisis.

A la consulta de Foniatría de Villa Clara asiste un alto por ciento de pacientes con afecciones vocales que son maestros, profesores, educadores, auxiliares pedagógicos, locutores y actores, etc. Si bien todos estos casos siguen tratamiento, no se había contado hasta el presente con una base de datos que permita almacenar las grabaciones, diagnósticos y demás de los pacientes. El desarrollo de una base de datos que reúna los informes de las consultas y las grabaciones realizadas debe funcionar como los cimientos para futuras investigaciones.

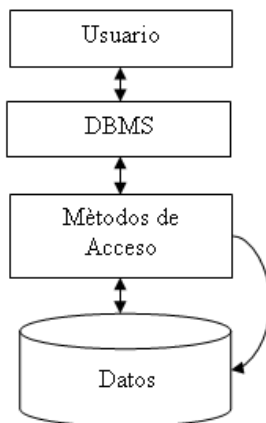
### 1.3 Base de datos

Este epígrafe aborda el tema de las bases de datos sus principales características así como funciones, además se refiere de manera específica a las bases de datos de voces fonéticas así como de voces patológicas.

Una base de datos es un repositorio donde se guarda información integrada que podemos almacenar y recuperar, o un conjunto de información almacenada en memoria auxiliar que permite acceso directo a un conjunto de programas que manipulan esos datos, a continuación se observan sus principales partes en Figura 4 [13].

Los componentes de una Base de Datos son:

- Hardware: constituido por un dispositivo de almacenamiento.
- Software: Sistema Administrador de Base de Datos (DBSM, por sus siglas en Inglés).
- Datos: los cuales están almacenados de acuerdo a la estructura externa y van a ser procesados para convertirse en información.



**Figura 4. Componentes de una Base de Datos.**

Las bases de datos son utilizadas por disimiles tipos de usuarios como se refiere a continuación:

- Usuario Final: persona que utiliza los datos, esta ve datos convertidos en información:
- Desarrollador de Aplicaciones: persona que desarrolla los sistemas que interactúan con la Base de Datos.
- Administrador de Base de Datos (DBA): persona que asegura integridad, consistencia, redundancia, seguridad, este es quien se encarga de realizar el mantenimiento diario o periódico de los datos.

Las personas que tienen acceso al sistema de mantenimiento de base de datos (DBMS) se clasifican de la siguiente manera:

- Usuarios ingenuos: aquellos que interactúan con el sistema por medio de aplicaciones permanentes.
- Usuarios sofisticados: aquellos con la capacidad de acceder a la información por medios de lenguajes de consulta.
- Programadores de aplicación: aquellos con un amplio dominio del tema, los cuales son capaces de generar nuevos módulos o utilerías capaces de manejar nuevos datos en el sistema.
- Usuarios especializados: aquellos que desarrollan módulos que no se refieren precisamente al manejo de los datos, sino a aplicaciones avanzadas como sistemas expertos, reconocimientos de imágenes, procesamiento de audio y demás [13].

### **1.3.1. Base de datos fonética, Modelo 4332.**

Esta base de datos fonética ilustra los sonidos, símbolos y convenciones de transcripciones para varios idiomas. En la mayoría de los casos, los artículos presentados tienen el propósito de ilustrar el grado de contraste fonológico, para representar una amplia gama de realizaciones fonéticas, y para dar ejemplos de funcionamiento de discurso [14].



En la actualidad, la base de datos está compuesta por más de 4.000 archivos de digitalizados separados en palabras u oraciones en 45 idiomas diferentes. Un número sustancial de las muestras digitalizadas tienen transcripciones AFI (Alfabeto Fonético Internacional) [15] sincronizadas a las formas de onda digitalizadas. Hay archivos adicionales con expresiones onomatopéyicas. Otros idiomas están en proceso de ser compilados para futuras actualizaciones de la base de datos.

Esta incluye un programa el cual proporciona acceso directo al menú de la base de datos. Este es un programa personalizado que utiliza herramientas del programa de Multi-Speech para explorar a fondo las propiedades acústicas. También ofrece la capacidad de agregar la transcripción para los datos [14].

#### **1.3.1.1 Manual**

El manual de 700 páginas de la base de datos incluye un inventario fonético del sistema de sonido de cada idioma, nombre de archivo, la transcripción fonológica y fonética, ortografía, y brillo además nos garantiza que el idioma oficial sea el inglés para todos los 4048 archivos.

#### **1.3.2 Base de datos de voces patológicas y Programa Modelo 4337.,**

La base de datos de trastornos de voz y su programa incluyen una base de datos de archivos y un programa para su visualización. Los espectrogramas de la vocal sostenida /a/ se presentan junto con un gráfico de MDVP (The Multi-Dimensional Voice Program) de voz.

La base de datos incluye más de 1.400 muestras de voces de vocales sostenidas y los primeros 12 segundos del *“Rainbow Passage”* de aproximadamente 700 sujetos. Esta base de datos se ha desarrollado como una ayuda en el análisis acústico y perceptivo de voces patológicas, ya sea para aplicaciones clínicas o de investigación. La base de datos incluye muestras de pacientes con una amplia variedad de patologías de tipo orgánico, neurológico, traumático, psicógena, y otros trastornos de la voz [14].

### 1.3.2.1 La base de datos

Massachusetts Eye and EarInfirmary (MEEI) ofrece un excelente modelo para la organización de la información del paciente y el examen. KayPENTAX ha seleccionado una gran colección representativa de voces y las examina con un conjunto completo de información sobre el paciente, resultados MDVP, muestras acústicas y el diagnóstico de los profesionales de MEEI. Todos los archivos de las señales y la información clínica para las voces patológicas se recogieron y se estructuraron en MEEI en el laboratorio de voces y habla. Prácticamente todos los pacientes tenían un extenso análisis de la función vocal incluyendo estroboscopia además del análisis aerodinámico y análisis acústico. Estas muestras con los datos y las evaluaciones asociadas representan muchos años de trabajo. A continuación se muestra en la Figura 5 la ventana con la que trabajan los especialistas a la hora de analizar los archivos de audio de un paciente [14].

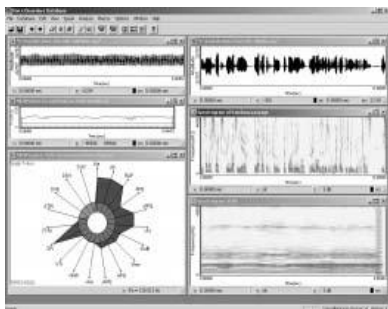


Figura 5. Ventana de trabajo para el análisis de archivos de audio.

### 1.3.2.2 Características de la base de datos de voces patológicas, modelo 4337.

- Presenta un CD-ROM con 1.400 archivos
- Amplia base de datos de información de los pacientes de voz que representa años de trabajo.
- Facilita la comprensión de los parámetros MDVP.
- El análisis en profundidad de los resultados del paciente utilizando tres parámetros objetivos.

- Los datos son recogidos y analizados por profesionales clínicos de habla y voz.

### **1.3.3 Aplicaciones**

Por primera vez, los hospitales y las universidades pueden tener acceso rápido e instantáneo, en un práctico formato de CD-ROM, a una amplia base de datos recolectados a través de un protocolo riguroso y técnicas directas de grabación digital. Todos los archivos acústicos se pueden cargar directamente en todos los programas de Multi-Speech y CSL. El CD-ROM incluye la información completa en archivos de los pacientes y MDVP, los resultados del análisis se obtienen en formato FoxPro es decir con una plantilla de visualización, también en formato de Excel o como archivos ASCII.

Anteriormente hemos reflejado la importancia de contar con una base de datos en nuestras consultas ya que nos ayuda a mejorar nuestro trabajo investigativo pero si queremos alcanzar resultados debemos crear software para el análisis de dichas bases, a continuación abordaremos este tema.

### **1.4. Software para el análisis de las señales de voz dentro de bases de datos.**

En este epígrafe se tratara lo relacionado con las herramientas de software disponibles para el análisis y desarrollo de las señales de voz y del habla en general.

#### **1.4.1. Praat**

Praat (*“hablar”* en idioma Holandés) es un software gratuito para el análisis y síntesis de voz escrito por Paul Boersma y David Weenink en el Departamento de Fonética de la Universidad de Ámsterdam. Incluye instalaciones para anotar datos de voz de formas de onda sincronizada que incluyan muestra de análisis espectrográficos, así como rutinas de procesamiento digital de voz. Praat ofrece el uso de varios niveles para la transcripción lineal. Los datos de voz anotada es la

base para las consultas como el número de artículos o mediciones de duración. Sin ningún conocimiento técnico Praat es fácil de usar [16].

#### 1.4.2. EMU

Al igual que Praat, EMU incluye instalaciones para anotar datos de voz, así como una biblioteca de rutinas para el procesamiento digital del habla. Se diferencia de Praat, sin embargo, al permitir anotaciones jerárquicas y autosegmentales y en la prestación de un lenguaje de consulta especial para la extracción de anotaciones de estas estructuras. Las anotaciones jerárquicas son más adecuadas para investigaciones lingüísticas que lineales. De esa manera los análisis son más complejos pues incluyen frases, palabras y estructuras silábicas.

El sistema de base de datos EMU es un conjunto de herramientas de software desarrollado para apoyar la investigación en fonética acústica y otras investigaciones.

Este ha sido desarrollado por McVeigh y Harrington en el 1992, Harrington, Cassidy, Fletcher y McVeigh en el 1993, así como Cassidy en el 1999, otra vez Cassidy y Harrington en 1996 y 2001 respectivamente y por ultimo Harrington, Cassidy, John y Scheffers en 2003, a su vez EMU es un conjunto integrado de herramientas para crear, anotar, consultar y analizar el habla. Desde 1988, cuando fue desarrollado por primera vez como fonética acústica en CSTR, la Universidad de Edimburgo (Taylor, Caley, Negro y Rey.,1999), el EMU ha proporcionado una interfaz transparente para un poderoso lenguaje gráfico y estadístico de programación (<biblio>). Otra característica distintiva del EMU es la organización de una base de datos de voz en una plantilla llamada archivo, que especifica la ubicación física de los archivos y los tipos de jerarquía así como de segmentos relacionados entre conjuntos de etiquetas [17].

El archivo de plantilla estaba incluido para facilitar el intercambio del archivo de voz a través de diferentes usuarios, trabajando en diferentes sitios, y posiblemente en diferentes plataformas. Otra característica, única de EMU, es que nos brinda la posibilidad de construir diversos tipos de estructuras de etiquetado

automáticamente, utilizando el script Tcl idioma [18]. Además el EMU utiliza como formato fundamental para sus archivos el SSFF. Recientemente, una interfaz como Praat ha sido construida, un sistema computacional para hacer la fonética [16], permitiendo que estructuras de etiquetado Praat se conviertan en EMU y viceversa [17].

#### **1.4.2.1 Estructura actual del EMU**

En los últimos tres años, el EMU 2.0 se ha provisto de una nueva estructura interna que permite al individuo modificar partes del sistema y que estos sean probados de forma independiente. Tanto el front-end y el back-end lo tienen diseñado de forma modular y este se puede ampliar sin mucho esfuerzo.

En versiones anteriores, el EMU consistía en tres partes las cuales proporcionan:

1-Funcionalidad de bases de datos.

2-Procesamiento de señales.

3- Presentación de la señales.

Estas piezas se integraron en el intérprete Tcl el cual es diferente dependiendo de las plataformas.

## **Capítulo 2. Materiales y Métodos:**

En este capítulo se describen las particularidades del software utilizado para la realización de este trabajo así como la descripción de los materiales y métodos empleados en el mismo como por ejemplo servidor de SQL, Microsoft Visual Studio y un NET Framework 4...

### **2.1. Diseño de la Base de Datos:**

Para organizar el flujo de información generado por los pacientes en la Consulta de Logopedia y Foniatria; se decide diseñar una base de datos a partir de los campos de las tablas de la Base de datos de la Kay-Elmetrics [9] y las sugerencias aportadas por los especialistas de la consulta provincial del mencionado servicio.

#### **2.1.1. Situación problemática:**

En la Consulta de Logopedia y Foniatria se atienden a un grupo de pacientes con diferentes patologías de Habla y voz. De ellos se desea conocer los siguientes datos:

Nombre completo, Edad, sexo, número de Historia Clínica, Habito toxico, labor que realiza, antecedentes patológicos personales (App), origen, Raza.

Además del proceso de diagnostico y tratamiento se desea conservar los siguientes datos:

Diagnóstico, Laringoscopia indirecta (descripción del resultado), predominio del timbre. También las grabaciones realizadas clasificadas de acuerdo a su tipo.

La situación planteada presenta semejanza con los problemas estudiados en la asignatura *Informática Médica* (4to año 1er semestre). Por lo cual selecciona la metodología propuesta (guía metodologica para diseñadores expertos) [19] para dar solución al problema. La guía consta de cinco pasos que se describen a continuación:

1- Candidatos a entidades:

Paciente: posee diagnóstico (s) y grabaciones (es).

Grabación: es generada por un Paciente.

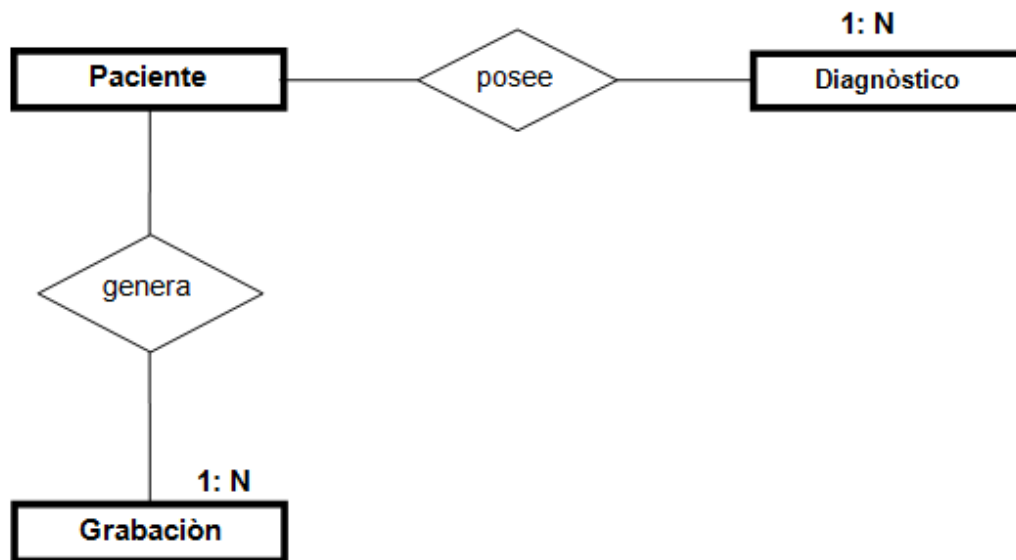
Diagnóstico: es emitido para un Paciente.

2- Representar todas las interrelaciones junto con su tipo de correspondencia.

Cada Paciente puede tener varios Diagnósticos (los pacientes presentan cambios a lo largo del tratamiento que conviene ser registrados; en este caso como Diagnósticos).

Cada Paciente puede tener varias Grabaciones.

3- Versión preliminar del diagrama Entidad/Relación Figura 6.



**Figura 6. Diagrama Entidad/Relación para la base de Datos de Pacientes.**

Los pasos 4 (Análisis de cardinalidades mínimas) y 5 (Análisis de redundancia) se omiten debido a la sencillez del diseño propuesto en 3.

Llegado a este punto se puede proceder a implementar la base de datos y la interfaz gráfica para el manejo de la misma. Para la realización de la Base de Datos se utiliza el language estandar SQL y para el desarrollo de la aplicación el Language C Sharp (C#) en el entorno de desarrollo integrado Microsoft Visual Studio. Estos lenguajes fueron seleccionados por haber sido impartidos en asignaturas de la carrera (Programación I y II).

Una descripción de estos lenguajes se realiza en los epígrafes siguientes.

## **2.2. Diseño del Software**

Para el diseño de un software se hace necesario el seguimiento de una metodología, esta se aplica actualmente en las casas desarrolladoras de software y en la Universidad de Ciencias Informáticas (UCI) [20] e incluye los siguientes pasos, que se amplían a continuación:

- 1.Nombrar el Software.
- 2.Objetivos.
- 3.Caracterización.
- 4.Información.
- 5.Diagrama de flujo.
- 6.Diseño general de las pantallas.
- 7.Diseño de la ayuda.
- 8.Requisitos funcionales.
- 9.Orientaciones.
- 10.Diseño.
- 11.Evaluación.

El método de recogida de información inicial empleado es entrevista a los especialistas de la consulta. A estos se les pregunta qué información desean guardar de los pacientes. A lo cual responden refiriéndose a la base de datos de la Kay-Elementrics [9] y le añaden una serie de características particulares del trabajo



que realizan en la consulta. Además de un fuerte énfasis en la necesidad de que la aplicación cree, almacene y reproduzca archivos de audio.

De esta entrevista surge el nombre del software (Pafcont) y se aclaran los objetivos del mismo:

Almacenar de manera organizada los datos de los pacientes, las grabaciones realizadas a los mismos y los diagnósticos emitidos en consulta. Realizar búsquedas de pacientes y grabaciones en la base de datos.

El software se caracteriza de la siguiente manera:

Temática: manejo de base de datos.

1. Nivel de los usuarios:

Especialistas clínicos del área de Logopedia y Foniatría.

2. Conocimientos propedéuticos:

Uso del mouse, el teclado y el sistema operativo Windows. Manejo básico de archivos de audio (crear, guardar, reproducir, eliminar).

Información requerida: Solo la información básica de manejo de base de datos y las variables a almacenar y buscar.

La idea de cómo debe ser la interfaz gráfica se plasma en un Diagrama de flujo Figura 7. En la misma se describe la manera en que la ventana principal se divide en dos partes, Lista de Pacientes y Paciente Seleccionado. Además se describe la arquitectura de los menús. El criterio a seguir se orienta hacia un estilo minimalista.

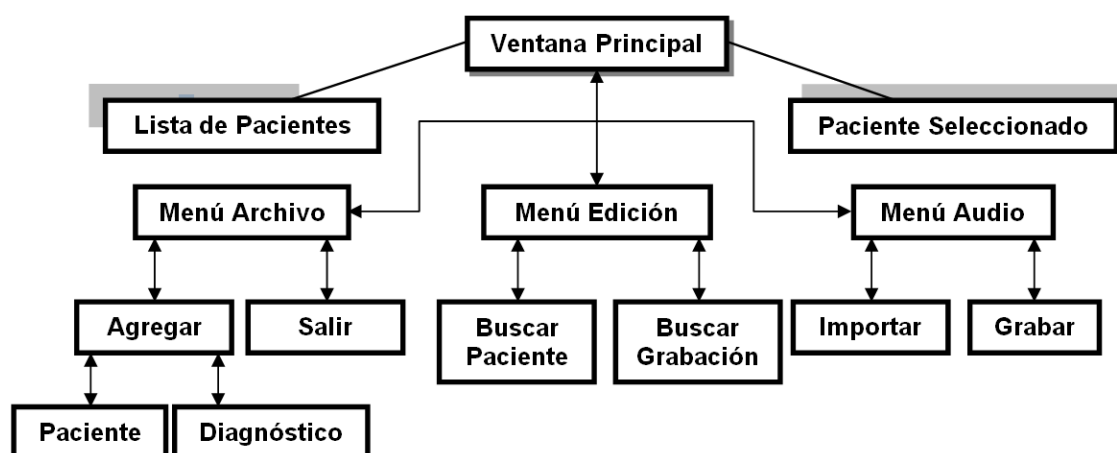


Figura 7 Diagrama de flujo de la estructura del software.

Diseño general de la pantalla Principal se describe en la Figura 8. Los diseños de las pantallas asociadas siguen este patrón con la ausencia de la Línea de Menú y con diferentes distribuciones de los paneles.

Barra de Título	
Línea de Menú	
Panel para Lista de Pacientes	Panel para Paciente Seleccionado

**Figura 8 Pantalla Principal**

El diseño de la ayuda en este caso debe ser un proceso en paralelo al trabajo de diploma, el cual incluye el aporte de los especialistas. Por este motivo los archivos de ayuda quedan como posible recomendación del trabajo de diploma.

Los requisitos funcionales del software propuesto se pueden resumir en los siguientes puntos:

- Visualizar información.
- Manipular archivos de audio
- Manipular la tarjeta de audio (grabación)
- Manipular la base de datos creada

Los Requisitos no funcionales no se han explorado a profundidad, pero pueden ser cumplidos por cualquier computadora de los últimos 10 años (debido a la baja complejidad computacional) que posea tarjeta de audio y sistema operativo Windows XP o posterior.

Las orientaciones de esta aplicación serian como herramienta de almacenamiento y gestión de información de los pacientes.

Como se plantea previamente el diseño predominante es minimalista

Por último la evaluación del mismo será a través de los criterios de los usuarios, los cuales deben ser la base para una versión 2.0 de Pafcont.

### 2.3. El lenguaje SQL

El SQL es el lenguaje estándar ANSI/ISO de definición, manipulación y control de bases de datos relacionales. Es un lenguaje declarativo: sólo hay que indicar qué se quiere hacer. En cambio, en los lenguajes procedimentales es necesario especificar cómo hay que hacer cualquier acción sobre la base de datos. El SQL es un lenguaje muy parecido al lenguaje natural; concretamente, se parece al inglés, y es muy expresivo. Por estas razones, y como lenguaje estándar, el SQL es un lenguaje con el que se puede acceder a todos los sistemas relacionales comerciales [21].

Empezamos con una breve explicación de la forma en que el SQL ha llegado a ser el lenguaje estándar de las bases de datos relacionales:

Al principio de los años setenta, los laboratorios de investigación Santa Teresa de IBM empezaron a trabajar en el proyecto System R. El objetivo de este proyecto era implementar un prototipo de SGBD relacional; por lo tanto, también necesitaban investigar en el campo de los lenguajes de bases de datos relacionales. A mediados de los años setenta, el proyecto de IBM dio como resultado un primer lenguaje denominado SEQUEL (*Structured English Query Language*), que por razones legales se denominó más adelante SQL (*Structured Query Language*). Al final de la década de los setenta y al principio de la de los ochenta, una vez finalizado el proyecto System R, IBM y otras empresas empezaron a utilizar el SQL en sus SGBD relacionales, con lo que este lenguaje adquirió una gran popularidad.

En 1982, ANSI (*American National Standards Institute*) encargó a uno de sus comités(X3H2) la definición de un lenguaje de bases de datos relacionales. Este comité, después de evaluar diferentes lenguajes, y ante la aceptación comercial del SQL, eligió un lenguaje estándar que estaba basado en éste prácticamente en su totalidad. El SQL se convirtió oficialmente en el lenguaje estándar de ANSI en el año 1986, y de ISO (*International Standards Organization*) en 1987. También ha sido adoptado como lenguaje estándar por FIPS (*Federal Information Processing Standard*), Unix X/Open y SAA (*Systems Application Architecture*) de IBM [21].

En el año 1989, el estándar fue objeto de una revisión y una ampliación que dieron lugar al lenguaje que se conoce con el nombre de SQL1 o SQL89. En el año 1992 el estándar volvió a ser revisado y ampliado considerablemente para cubrir carencias de la versión anterior. Esta nueva versión del SQL, que se conoce con el nombre de SQL2 o SQL92.

Como veremos más adelante, aunque aparezca sólo la sigla SQL, siempre nos estaremos refiriendo al SQL92, ya que éste tiene como subconjunto el SQL89, por lo tanto, todo lo que era válido en el caso del SQL89 lo continuará siendo en el SQL92.

De hecho, se pueden distinguir tres niveles dentro del SQL92:

- 1) El nivel introductorio (*entry*), que incluye el SQL89 y las definiciones de *clave primaria* y *clave foránea* al crear una tabla.
- 2) El nivel intermedio (*intermediate*), que, además del SQL89, añade algunas ampliaciones del SQL92.
- 3) El nivel completo (*full*), que ya tiene todas las ampliaciones del SQL92.

## 2.4. Net Framework

**NET** es un framework de Microsoft que hace un énfasis en la transparencia de redes, con independencia de plataforma de hardware y que permita un rápido desarrollo de aplicaciones. Basado en ella, la empresa intenta desarrollar una estrategia horizontal que integre todos sus productos, desde el sistema operativo hasta las herramientas de mercado. NET podría considerarse una respuesta de Microsoft al creciente mercado de los negocios en entornos Web, como competencia a la plataforma Java de Oracle Corporation y a los diversos framework de desarrollo web basados en PHP. Su propuesta es ofrecer una manera rápida y económica, a la vez que segura y robusta, de desarrollar aplicaciones –o como la misma plataforma las denomina, soluciones– permitiendo una integración más rápida y ágil entre empresas y un acceso más simple y universal a todo tipo de información desde cualquier tipo de dispositivo [22].

La plataforma .NET de Microsoft es un componente de software que puede ser añadido al sistema operativo Windows. Provee un extenso conjunto de soluciones predefinidas para necesidades generales de la programación de aplicaciones, y administra la ejecución de los programas escritos específicamente con la plataforma. Esta solución es el producto principal en la oferta de Microsoft, y pretende ser utilizada por la mayoría de las aplicaciones creadas para la plataforma Windows. .NET Framework se incluye en Windows Server 2008, Windows Vista y Windows 7. De igual manera, la versión actual de dicho componente puede ser instalada en Windows XP, y en la familia de sistemas operativos Windows Server 2003. Una versión "reducida" de .NET Framework está disponible para la plataforma Windows Mobile, incluyendo teléfonos inteligentes.

#### **2.4.1 Estructura interna del entorno de ejecución en lenguaje común.**

El CLR es el verdadero núcleo del framework de .NET, entorno de ejecución en el que se cargan las aplicaciones desarrolladas en los distintos lenguajes, ampliando el conjunto de servicios del sistema operativo (W2k y W2003). Permite integrar proyectos en distintos lenguajes soportados por la plataforma .Net, como C++, Visual Basic, C#, entre otros [22].

La herramienta de desarrollo compila el código fuente de cualquiera de los lenguajes soportados por .NET en un código intermedio, el CIL (*Common Intermediate Language*) antes conocido como MSIL (*Microsoft Intermediate Language*), similar al código de byte de Java. Para generarlo, el compilador se basa en la especificación CLS (*Common Language Specification*) que determina las reglas necesarias para crear el código MSIL compatible con el CLR.

Para ejecutarse se necesita un segundo paso, un compilador JIT (*Just-In-Time*) es el que genera el código máquina real que se ejecuta en la plataforma del cliente. De esta forma se consigue con .NET independencia de la plataforma de hardware. La compilación JIT la realiza el CLR a medida que el programa invoca métodos. El código ejecutable obtenido se almacena en la memoria caché del ordenador,

siendo recompilado de nuevo sólo en el caso de producirse algún cambio en el código fuente [22].

#### **2.4.2 Características**

Es el encargado de proveer lo que se llama código administrado, es decir, un entorno que provee servicios automáticos al código que se ejecuta. Los servicios son variados:

- Cargador de clases: permite cargar en memoria las clases.
- Compilador MSIL a nativo: transforma código intermedio de alto nivel independiente del hardware que lo ejecuta a código de máquina propio del dispositivo que lo ejecuta.
- Administrador de código: coordina toda la operación de los distintos subsistemas del corredor de tiempo para los lenguajes frecuentes.
- Recolector de basura: elimina de memoria objetos no utilizados automáticamente.
- Motor de seguridad: administra la seguridad del código que se ejecuta.
- Motor de depuración: permite hacer un seguimiento de la ejecución del código aun cuando se utilicen lenguajes distintos.
- Verificador de tipos: controla que las variables de la aplicación usen el área de memoria que tienen asignado.

#### **2.5. Microsoft Visual Studio**

Este es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) para sistemas operativos Windows. Soporta varios lenguajes de programación, tales como Visual C++, Visual C# utilizado para la programación de este trabajo, Visual J#, y Visual Basic.

Visual Studio permite a los desarrolladores crear sitios y aplicaciones web, así como servicios web en cualquier entorno que soporte la plataforma .NET a partir de la versión .NET 2002. Así se pueden crear aplicaciones que se comuniquen entre estaciones de trabajo, páginas web y dispositivos móviles [23].

### **2.5.1. Versiones**

A partir de la versión 2005 Microsoft ofrece gratuitamente las ediciones *Express*, que son varias ediciones básicas separadas por lenguajes de programación o plataforma enfocadas para novatos y entusiastas. Estas ediciones son iguales al entorno de desarrollo comercial pero sin características avanzadas. Dichas ediciones son:

- Visual Basic Express Edition
- Visual C# Express Edition
- Visual C++ Express Edition
- Visual J# Express Edition (Desapareció en Visual Studio 2008)
- Visual Web Developer Express Edition (para programar en ASP.NET)
- Visual F# (Apareció en Visual Studio 2010, es parecido al J#)\*

Adicionalmente, Microsoft ha puesto gratuitamente a disposición de todo el mundo una versión reducida de MS SQL Server llamada SQL Server Express Edition cuyas principales limitaciones son que no soporta bases de datos superiores a 4 GB de tamaño, únicamente se ejecuta en un procesador y emplea 1 GB de RAM como máximo, y no cuenta con el Agente de SQL Server [23].

#### **2.5.1.1 Visual Studio 6.0**

Se lanzó en 1998 y fue la última versión en ejecutarse en la plataforma Win9x. Los números de versión de todas las partes constituyentes pasaron a 6.0, incluyendo Visual J++ y Visual InterDev, que se encontraban en las versiones 1.1 y 1.0 respectivamente. Esta versión fue la base para el sistema de desarrollo de

Microsoft para los siguientes 4 años, en los que Microsoft migró su estrategia de desarrollo al .NET Framework [23].

### **2.5.1.2 Visual Studio .NET (2002)**

En esta versión se produjo un cambio sustancial, puesto que supuso la introducción de la plataforma .NET de Microsoft. .NET es una plataforma de ejecución intermedia multilenguaje, de forma que los programas desarrollados en .NET no se compilan en lenguaje máquina, sino en un lenguaje intermedio (CIL - Common Intermediate Language) denominado Microsoft Intermediate Language (MSIL) [24].

Visual Studio .NET 2002 supuso también la introducción del lenguaje C Sharp (C#), un lenguaje nuevo diseñado específicamente para la plataforma .NET, basado en C++ y Java. El lenguaje Visual Basic fue remodelado completamente y evolucionó para adaptarse a las nuevas características de la plataforma .NET, haciéndolo mucho más versátil y dotándolo con muchas características de las que carecía.

### **2.5.1.3 Visual Studio .NET 2003**

Visual Studio .NET 2003 supone una actualización *menor* de Visual Studio .NET. Se actualiza el .NET Framework a la versión 1.1. También se añade soporte con el fin de escribir aplicaciones para determinados dispositivos móviles, ya sea con ASP.NET o con el .NET Compact Framework. Además, el compilador de Visual C++ se mejora para cumplir con más estándares: el Visual C++ Toolkit 2003.

Visual Studio 2003 se lanza en 4 ediciones: Académica, Profesional, Enterprise Developer y Arquitectura Enterprise. La edición Arquitectura Enterprise incluía una implementación de la tecnología de modelado Microsoft Visio, que se centraba en la creación de representaciones visuales de la arquitectura de la aplicación basadas en UML la versión interna de Visual Studio .NET 2003 es la 7.1 aunque el formato del archivo que emplea es el 8.0 [24].



#### **2.5.1.4 Visual Studio 2005**

Visual Studio 2005 se empezó a comercializar a través de Internet a partir del 4 de octubre de 2005 y llegó a los comercios a finales del mes de octubre en inglés. En castellano no salió hasta el 4 de febrero de 2006. Microsoft eliminó la coletilla *.NET* de su nombre, pero eso no indica que se alejara de la plataforma *.NET*, de la cual se incluyó la versión 2.0.

Visual Studio 2005 tiene varias ediciones radicalmente distintas entre sí: Express, Standard, Professional, herramientas para Office y 5 ediciones Visual Studio para sistema en equipo. Éstas últimas se proporcionan conjuntamente con suscripciones a MSDN cubriendo los 4 principales roles de la programación: Arquitectura, Software Developers, Testers y Profesionales Base de Datos. La funcionalidad combinada de las 4 ediciones como sistema de equipo se ofrecía como la edición suite equipo. Por otra parte, el sistema de herramientas para el Microsoft Office está diseñada para extender su funcionalidad [24].

#### **2.5.1.5 Visual Studio 2008**

Visual Studio 2008 fue publicado (RTM) el 17 de noviembre de 2007 en inglés, mientras que la versión en castellano no fue publicada hasta el 2 de febrero de 2008 [25].

A las mejoras de desempeño, escalabilidad y seguridad con respecto a la versión anterior, se agregan, entre otras, las siguientes novedades:

- Visual Studio 2008 permite incorporar características del nuevo Windows Presentation Foundation sin dificultad tanto en los formularios de Windows existentes como en los nuevos. Ahora es posible actualizar el estilo visual de las aplicaciones al de Windows Vista debido a las mejoras en Microsoft Foundation Class Library (MFC) y Visual C++.

- Visual Studio 2008 permite mejorar la interoperabilidad entre código nativo y código manejado por .NET. Esta integración más profunda simplificará el trabajo de diseño y codificación.

#### **2.5.1.6 Visual Studio 2010**

Visual Studio 2010 es la versión más reciente de esta herramienta, acompañada por .NET Framework 4.0. La fecha del lanzamiento de la versión final fue el 12 de abril de 2010 [23].

Hasta ahora, uno de los mayores logros de la versión 2010 de Visual Studio ha sido el de incluir las herramientas para desarrollo de aplicaciones para Windows 7.

Además ofrece la posibilidad de crear aplicaciones para muchas plataformas de Microsoft, como Windows, Azure, Windows Phone 7 o Sharepoint. Microsoft ha sido sensible a la nueva tendencia de las pantallas táctiles y con este Visual Studio 2010 también es posible desarrollar aplicativos para pantallas multitáctiles [26].

A continuación se hablara del lenguaje de programación utilizado en este trabajo que en este caso fue el C Sharp(C#), además se abordaran sus principales características así como algo del origen del mismo

### **2.6. C Sharp (C#)**

#### **2.6.1. Origen**

Durante el desarrollo de la plataforma .NET, las bibliotecas de clases fueron escritas originalmente usando un sistema de código gestionado llamado Simple Managed C (SMC). En enero de 1999, Anders Hejlsberg formó un equipo con la misión de desarrollar un nuevo lenguaje de programación llamado Cool (Lenguaje C orientado a objetos). Este nombre tuvo que ser cambiado debido a problemas de marca, pasando a llamarse C#. La biblioteca de clases de la plataforma .NET fue migrada entonces al nuevo lenguaje.

Hejlsberg lideró el proyecto de desarrollo de C#. Anteriormente, ya había participado en el desarrollo de otros lenguajes como Turbo Pascal, J++ [27].

C# (pronunciado *si sharp* en inglés) es un lenguaje de programación orientado a objetos desarrollado y estandarizado por Microsoft como parte de su plataforma .NET, que después fue aprobado como un estándar por la ECMA (ECMA-334) e ISO (ISO/IEC 23270). C# es uno de los lenguajes de programación diseñados para la infraestructura de lenguaje común [27].

Su sintaxis básica deriva de C/C++ y utiliza el modelo de objetos de la plataforma .NET, similar al de Java, aunque incluye mejoras derivadas de otros lenguajes. El nombre C Sharp fue inspirado por la notación musical, donde '#' (sostenido, en inglés *sharp*) indica que la nota (C es la nota do en inglés) es un semitono más alta, sugiriendo que C# es superior a C/C++. Además, el signo '#' se compone de cuatro signos '+' pegados.

Aunque C# forma parte de la plataforma .NET, ésta es una API, mientras que C# es un lenguaje de programación independiente diseñado para generar programas sobre dicha plataforma. Ya existe un compilador implementado que provee el marco Mono - DotGNU, el cual genera programas para distintas plataformas como Windows, Unix, Android, iOS, Windows Phone, Mac OS y GNU/Linux [27].

#### **2.6.2. En la actualidad existen los siguientes compiladores para el lenguaje C#.**

- Microsoft .NET Framework 2.0 (SDK) incluye un compilador de C#, pero no un IDE.
- Microsoft Visual Studio, IDE por excelencia de este lenguaje.
- Sharp Develop, IDE libre para C# bajo licencia GNU LGPL, con una interfaz muy similar a Microsoft Visual Studio.
- Mono, es una implementación con licencia GNU GPL de todo el entorno .NET desarrollado por Novell. Como parte de esta implementación se incluye un compilador de C#.

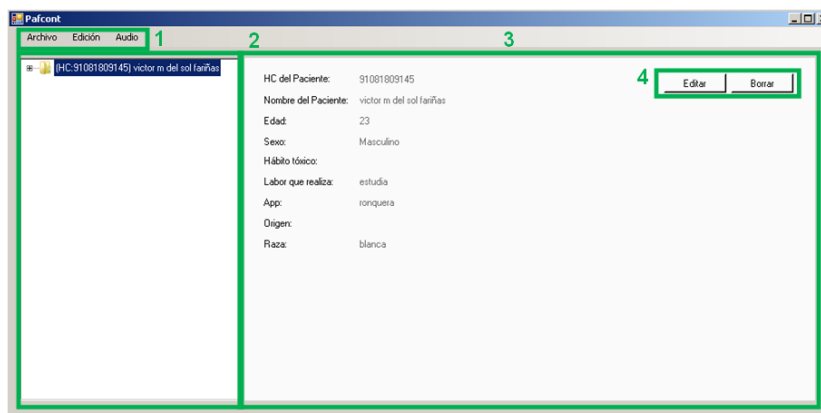
**2.6.3. El estándar ECMA-334 lista las siguientes metas en el diseño para C#:**

- Lenguaje de programación orientado a objetos simple, moderno y de propósito general.
- Capacidad para desarrollar componentes de software que se puedan usar en ambientes distribuidos.
- Portabilidad del código fuente.
- Fácil migración del programador al nuevo lenguaje, especialmente para programadores familiarizados con C, C++ y Java.
- Soporte para internacionalización.
- Aplicaciones económicas en cuanto a memoria y procesado.

## CAPÍTULO 3: RESULTADOS Y DISCUSIÓN.

### 3.1. Software Pafcont.

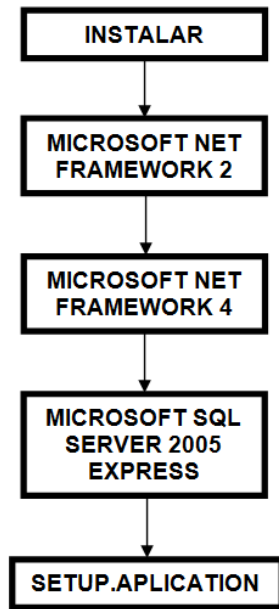
A partir de los materiales y métodos explicados el Capítulo 2, se implementó un software (Pafcont, nombre sugerido por los especialistas) cuya interfaz gráfica se muestra en la Figura 9. Este nos brinda diferentes facilidades que se describen a continuación. Además se ha incluido una breve descripción del metodo de empleo de la aplicación.



**Figura 9 Aplicación Pafcont (Interfaz gráfica).**1. Menús de la aplicación. 2. Lista de Pacientes. 3. Paciente seleccionado. 4. Controles para editar el Paciente seleccionado.

#### 3.1.1.Instalación de la Aplicación:

El presente trabajo se orienta a dar solución a un problema existente en la Consulta de Logopedia y Foniatría. Esta consulta cuenta con una computadora con las condiciones para utilizar un sistema operativo Windows XP, por ello la aplicación requiere ser operativa en dicho sistema. Luego de un análisis de las funciones que necesitaba el programa y las que podía soportar el sistema operativo se llega al siguiente diagrama en bloques que describe como realizar el proceso de instalación como se observa en la Figura 10.

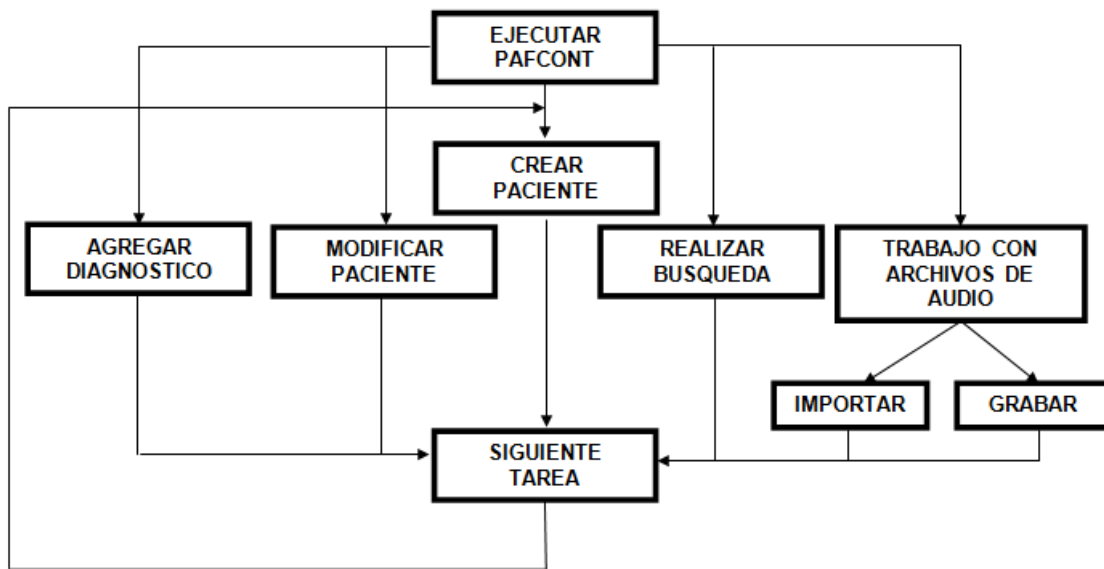


**Figura 10. Diagrama de Flujo del proceso de instalación.**

La versión de SQL utilizada es la más “ligera” disponible en la red de nuestro centro, pero requiere de MICROSOFT NET FRAMEWORK 2 y la aplicación de MICROSOFT NET FRAMEWORK 4 es requerida por el software en sí, lo cual incrementa en un paso el proceso de instalación. Esto se debe corregir en futuras versiones de Pafcont.

### **3.1.2. Modo de empleo de la aplicación**

Con el objetivo de facilitar el uso de la aplicación se presenta el siguiente diagrama de flujo Figura 11 que muestra las posibilidades que ofrece el software una vez ejecutado por el usuario.



**Figura 11 Diagrama de Flujo para el correcto uso de Pafcont.**

### 3.1.3. Manejo de la Aplicación.

A continuación se describen las cinco operaciones básicas que se pueden realizar en la aplicación:

#### 3.1.3.1. Introducir un nuevo paciente a la base de datos.

Para hacer este procedimiento se despliega el menú Archivo, el submenú Agregar y se pulsa el botón Paciente Figura 12. Luego aparece la ventana que se muestra en Figura 13 con el formulario que permite llenar todos los campos de la tabla Paciente.



**Figura 12 Agregar Paciente.**

The screenshot shows the 'Pafcont' application window with a menu bar containing 'Archivo', 'Edición', and 'Audio'. A sub-window titled 'Nuevo Paciente' is open, featuring a list of input fields on the left and two larger text areas on the right. The fields on the left are: 'No. Historia Clínica', 'Nombre', 'Primer Apellido', 'Segundo Apellido', 'Raza', 'Edad', 'Sexo' (a dropdown menu), 'Labor que realiza', and 'App'. The text areas on the right are labeled 'Hábito Tóxico' and 'Origen'. At the bottom right of the sub-window are 'Aceptar' and 'Cancelar' buttons. The main window also has 'Editar' and 'Borrar' buttons at the top right.

**Figura 13 Formulario para Agregar Paciente**

### 3.1.3.2. Ingresar un diagnóstico a un paciente previamente seleccionado.

En este caso es estrictamente necesario seleccionar al paciente antes de comenzar (de no existir pacientes emite un mensaje de advertencia), de otra manera el diagnóstico se agrega al paciente seleccionado anteriormente. Para esta tarea se despliega el menú Archivo, el submenú Agregar y se pulsa el botón Diagnóstico Figura 14. Instantáneamente se muestra la ventana con el formulario Figura 15 que permite llenar todos los campos de la tabla Diagnóstico.



**Figura 14. Agregar Diagnóstico.**



**Figura 15. Formulario para Agregar Diagnóstico.**

### 3.1.3.3. Modificar un paciente.

Es una posibilidad que ofrecen los botones señalados que se muestran en la Figura 16. En el caso del botón Editar abre el formulario mostrado en Figura 17; y en el caso de borrar, simplemente elimina el paciente, sus diagnósticos y grabaciones.

**Figura 16. Modificar Paciente. Enmarcados en rojo botones para modificar la tabla Paciente del paciente seleccionado.**

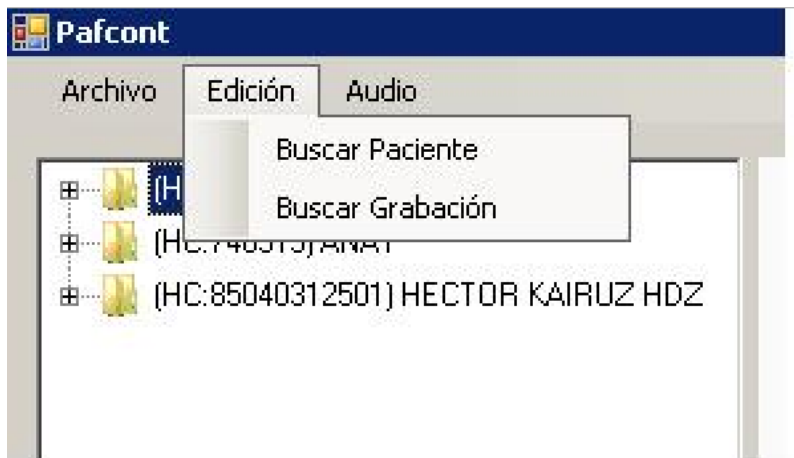
The screenshot shows a window titled 'Pafcont' with a menu bar containing 'Archivo', 'Edición', and 'Audio'. A modal dialog box titled 'Nuevo Paciente' is open. It contains the following fields and controls:

- No. Historia Clínica: 91081809145
- Nombre: victor manuel
- Primer Apellido: del sol
- Segundo Apellido: farías
- Raza: blanca
- Edad: 23
- Sexo: Masculino (selected from a dropdown)
- Labor que realiza: App
- Hábito Tóxico: (empty text box)
- Origen: (empty text box)
- Buttons: 'Aceptar' and 'Cancelar' at the bottom of the dialog.
- Buttons: 'Editar' and 'Borrar' on the right side of the main window.

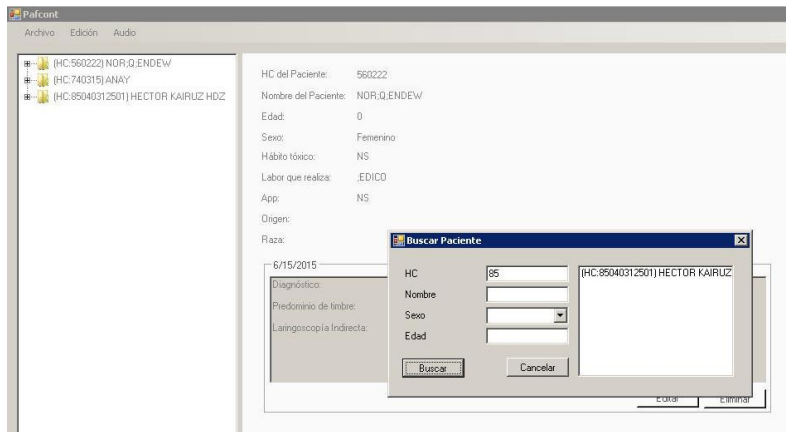
**Figura 17. Formulario para modificar Paciente.**

#### 3.1.3.4. Realizar Búsquedas.

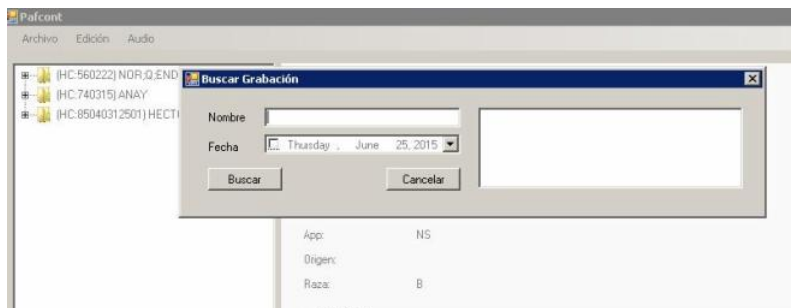
De acuerdo con las solicitudes de los especialistas se hace necesario realizar búsquedas de pacientes atendiendo diferentes atributos del mismo y de manera semejante para las grabaciones. Para ello se despliega el menú Edición Figura 18, y puede escoger entre Buscar Paciente o Buscar Grabación Figura 19 y Figura 20 respectivamente.



**Figura 18. Realizar Búsquedas**



**Figura 19. Ventana de Búsqueda de Paciente.**



**Figura 20. Ventana de Búsqueda de Grabación.**

### 3.1.3.5 Operaciones con Archivos de Audio.

Para realizar Operaciones con Archivos de Audio. Es necesario desplegar el menú Audio Figura 21 y escoger entre importar un archivo o grabarlo con la aplicación. En el caso de Importar es necesario que el archivo sea de extensión \*.WAV, codificación PCM, mono. Para ello se implementa un filtro en la ventana de Importar (esta ventana es una modificación de una plantilla de ventana del software de desarrollo) Figura 22 . Esta facilidad se incluye porque existe un elevado número de archivos de audio de pacientes que se deben incluir a la base de datos. Cuando se desea grabar con la aplicación aparece la ventana mostrada en Figura 23 . En este caso se graba a una frecuencia de muestreo de 44.1 KHz con 16bits, esta interfaz no ofrece posibilidad de configurar los mencionados parámetros debido a que se desea almacenar los datos con la mayor calidad posible.



Figura 21. Menú Audio.

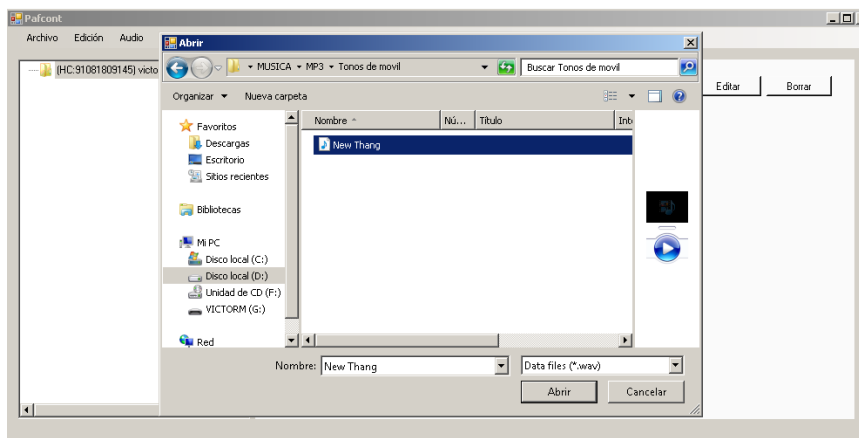


Figura 22. Ventana para Importar archivo de audio.

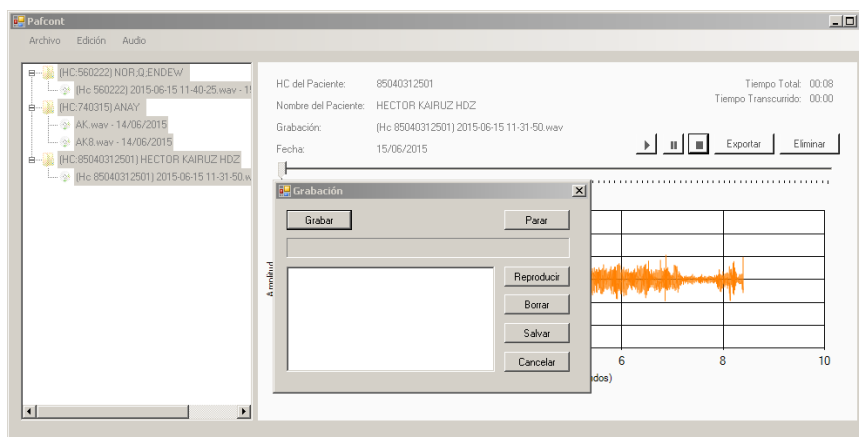
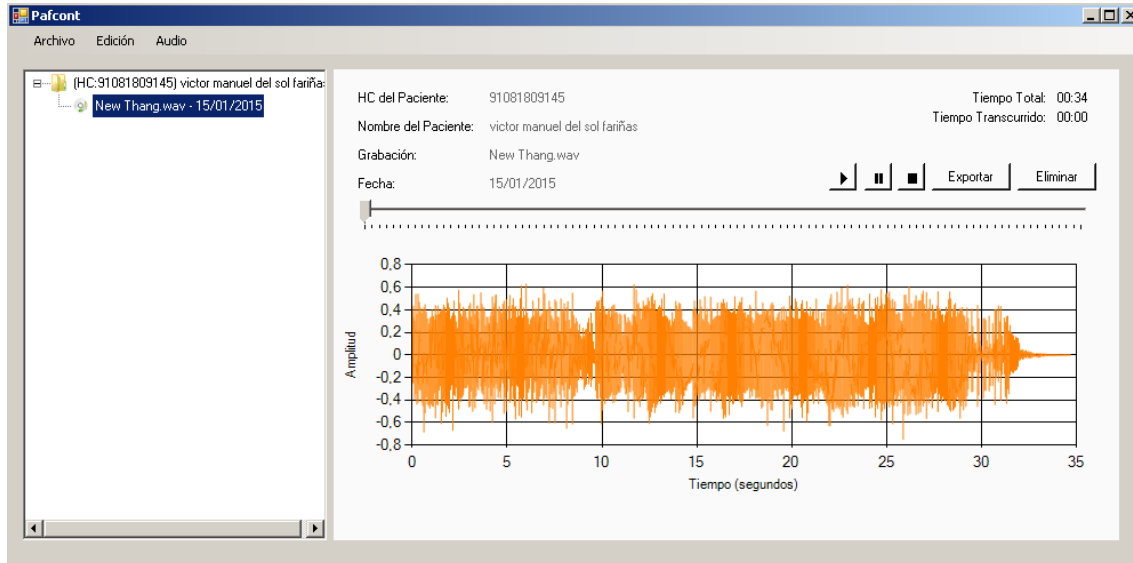


Figura 23. Ventana para Grabar.

Finalmente cuando un archivo de audio es incorporado a un paciente este se puede seleccionar y realizar diferentes operaciones con el mismo. Ejemplo de ello es la Figura 24 donde se observa que se puede reproducir, exportar o eliminar un archivo \*.WAV.



**Figura 24. Ejemplo de manejo de un archivo de audio.**

Otra de las demandas de los especialistas es sobre la posibilidad de tener un respaldo de la base de datos, este tema es como medida de prevención en caso de rotura de la computadora de la consulta. Para ello durante la instalación del software se establece una carpeta con localización en C:\Pafcont la cual contiene los archivos de la base de datos. Para resguardar los mismos es suficiente con copiarlos a otra unidad de almacenamiento y en caso de pérdidas de información o si se desea instalar la aplicación con la base de datos en otra computadora solo basta con reemplazar los archivos existentes en el mencionado directorio con aquellos que fueron copiados.

## Conclusiones Parciales

Como conclusiones parciales se puede decir que Pafcont es una interfaz gráfica sencilla y amena y que su diseño y funcionalidad fue aprobado con leves críticas por los especialistas.

## **CONCLUSIONES Y RECOMENDACIONES**

### **Conclusiones:**

1. Se diseñó de conjunto con los especialistas la base de datos y se colegiaron los requerimientos del software Pafcont.
2. Se implementa el software exitosamente.
3. Se crea una interfaz gráfica para el software, que ofrece al especialista un ambiente ameno y fácil para realizar el trabajo.
4. Se evalúa y aprueba por los especialistas el proyecto final.

## **Recomendaciones**

1. Crear el archivo de ayuda de la aplicación.
2. Modificar la búsqueda de pacientes (incluir búsqueda por profesión).
3. Incluir en versiones posteriores la posibilidad del cálculo de medidas objetivas y el espectrograma tiempo frecuencia para los archivos de audio.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] M. M. a. J. Schroeter, "Speech Production Models and Their Digital Implementations," in *Digital Signal Processing Handbook*, V. K. Madisetti and D. B. Williams,, 1999.
- [2] H. A. K. H. Díaz., "Detección y marca de los cambios espectrales notorios en voces patológicas," 2009.
- [3] M. Zañartú., "Aplicaciones del análisis acústico en los estudios de la voz humana," Seminario Internacional de Acústica," 2003.
- [4] M. J. L. A. a. J. González., "Medida de la inteligibilidad en el habla disártrica," *Revista de Logopedia, Foniatría y Audiología*," vol. 24, pp. 33-43, 2004.
- [5] H. Fletcher, J. B. Allen, and W. D. Ward, "ASA edition of speech and hearing in communication," *The Journal of the Acoustical Society of America*, vol. 100, pp. 685-685, 1996.
- [6] A. F. Espinoza, "Reconocimiento dn Palabras Aisladas en Castellano," ed: Inictel. Dirección de Investigación y Desarrollo, 1993.
- [7] I. Cobeta, F. Núñez, and S. Fernández, *Patología de la voz*: Marge books, 2013.
- [8] B. Boyanov and S. Hadjitodorov, "'Acoustic analysis of pathological voices:," vol. vol.16, pp. 74-82, 1997.
- [9] "Kay Elemetrics: Multi-dimensional voice program (MDVPTM), model 4305," 1993.
- [10] E. C. Guerra, "A modern approach to dysarthria classification," 2003.
- [11] J. H. L. Hansen, "A nonlinear operator-based speech feature analysis method with application to vocal fold pathology assessment, *IEEE Transactions on Biomedical Engineering*," pp. 300-312, 1998.
- [12] J. G. P. J. R. Deller, J. H. L. Hansen, " Discrete Time Processing of Speech Signals," 1993.
- [13] C. J. Date, *Introducción a los sistemas de bases de datos*: Pearson Educación, 2001.
- [14] KayPENTAX, "A Division of PENTAX Medical Company."
- [15] J. I. Hualde, *The sounds of Spanish*: Cambridge University Press New York, 2005.
- [16] P. a. W. D. Boersma, "Praat, a system for



doing phonetics by computer," 2001.

- [17] C. S. Harrington J., Fletcher J. and McVeigh A., "The mu+ system for corpusbased speech

research, *Computer Speech and Language*," pp. 305-331, 1993.

- [18] B. B. Welch, K. Jones, and J. Hobbs, *Practical programming in Tcl and Tk* vol. 1: Prentice Hall Professional, 2003.

- [19] M. Castaño Adoración, *Diseño de bases de datos: problemas resueltos*, 2001.

- [20] M. Linares Río, "Diseño de un software para la enseñanza de la asignatura Programación y Gestores de Bases de Datos en la carrera de Tecnología de la Salud," *Revista de Ciencias Médicas de Pinar del Río*, vol. 18, pp. 841-851, 2014.

- [21] J. Melton, "Database language sql," in *Handbook on Architectures of Information Systems*, ed: Springer, 1998, pp. 103-128.

- [22] P#, "A concurrent Prolog for .NET."

- [23] "Introducing Visual Studio 97."

- [24] V. C. D. Center, "Frequently Asked Questions About Visual C# .NET " 2003.

- [25] "• Visual Studio Express Editions."

- [26] "Description of Visual Studio 2012 Update 3," 06-26 2013.

- [27] J. Kovacs, "C#/.NET HistoryLesson," 2007.