

UNIVERSIDAD POLITECNICA DE VALENCIA

ESCUELA POLITECNICA SUPERIOR DE GANDIA

Master en Ingeniería Acústica



UNIVERSIDAD
POLITECNICA
DE VALENCIA



ESCUELA POLITECNICA
SUPERIOR DE GANDIA

**“Desarrollo de herramienta software
para la realización de Test de audición
con ruido.”**

TRABAJO FINAL DE MASTER

Autor/a:

Josep Rodríguez Sendra

Tutor/a:

Miguel Ferrer Contreras

Gema Piñero Sipán

GANDIA, 2015

Resumen

Hearing In Noise Test (HINT), Prueba de audición con Ruido, es un test que se utiliza para calcular la inteligibilidad tanto en personas con deficiencias auditivas como en personas con la audición sana. Se ha emulado el test HINT, añadiendo distintas variaciones, ya que originalmente se utiliza con ruido blanco, En este TFM para conseguir mayor semejanza con la cotidianidad se han utilizado distintos ruidos (Carretera, Cafetería, Ruido aleatorio y Ruido de Estación de tren) para estimar la relación Señal a Ruido que proporciona un 50% de inteligibilidad. Las listas de palabras son los estímulos más utilizados en las audiometrías verbales por las ventajas que representan en la práctica clínica (son rápidas y es un método estándar internacional), pero también presentan ciertas desventajas entre las que cabe destacar, que no representan una situación de comunicación cotidiana puesto que en la vida real las conversaciones se dan empleando frases y/o discursos continuos y no palabras aisladas. Se han utilizado 50 frases y evaluado a 30 individuos sin deficiencias auditivas de edades entre 23 y 58 años, con el fin de conseguir un resultado promedio de la SNR ("Signal to Noise Ratio") buscada para cada uno de los distintos ruidos. Se ha realizado un programa ejecutable con *MATLAB*, para realizar este test, finalmente se han analizado los resultados los cuales nos servirán para investigaciones futuras centradas en la inteligibilidad de las personas.

Palabras clave: Inteligibilidad, Test de audición, audiometría, HINT, Percepción del habla, Mejora del habla, Calidad auditiva, Ayudas a la audición.

Abstract

Hearing In Noise Test (HINT) is a test that is used to calculate the intelligibility both hearing impaired and in persons with healthy hearing. Has emulated the test HINT, adding different variations, as originally used with white noise, in this TFM for greater resemblance to the everyday have been used various noises (road, Café, random noise and noise from train station) to calculate 50% intelligibility. The Word lists are best stimuli used in verbal audiometry by the advantages they represent in clinical practice (they are fast and is an international standard method) but also have certain disadvantages, among which include, without representing a situation of because everyday communication in real life conversations occur using phrases and / or continuing talks and not isolated words. 50 sentences were used and evaluated 30 individuals without hearing impairments with ages 23 and 58, in order to obtain an average result of the SNR ("Signal to Noise Ratio") for different noises. There has been an executable program with *MATLAB*, to perform this test, finally we analyzed the results of which will serve for future research focused on the intelligibility of the people.

Key words: Intelligibility, Hearing test, Audiometry, HINT, Speech perception, Speech Improvement, Quality hearing, Hearing aid.

Agradecimientos

Me gustaría Agradecer y dedicar este trabajo a toda la gente que ha estado a mi lado, apoyándome desde el primer segundo, En ocasiones me he puesto muy pesado y muy monotema, pero el resultado ha merecido la pena. Gracias al grupo Gtac del iTeam, ya que han sido muchas horas en el laboratorio con Juan, gracias por los momentos de comer, Christian, Pablo, Jorge, José Antonio Laura... esos momentos ayudan a desconectar y seguir con fuerza. Por supuesto muchas gracias a mi tutor Miguel Ferrer ya que siempre que he necesitado de tu ayuda has estado disponible y con mucha paciencia. Gracias a Gema Piñero por todo y a María de Diego.

Gracias a todos los profesores que me han ido enseñando durante el master y con quienes no me gustaría perder el contacto.

Gracias a mis compañeros de master, Alberto, Jorge Enrique, Elisa y Manuel.

Gracias a els meus Telecoms, Faus, Hector, Jorge, Marc, David, Nacho, Veiga, Peter y Adri.

Gracias a Caselles, Joan, Victor, Rixi, Carcaixent, Casildo, Bolu, Pau. Los amigos son importantes y no hay que descuidarlos, Gracias a Smoking Souls por todos los momentos.

Gracias a mi Madre Xelo que siempre está ahí para todo lo que haga falta y que ha hecho de mi lo que soy. Gracias a mis abuelos en especial a mi abuela Rosa que sé que desde ahí estarás conmigo.

Gracias a Rafa mi hermano.

Y sobre todo gracias a ti Elvira, todos mis pensamientos están contigo, este trabajo te lo dedico especialmente a ti, la que me ha estado aguantando hasta el último momento.

ÍNDICE

I.	Resumen-Abstract	
II.	Agradecimientos	
1.	Presentación y motivación.....	5
1.1.	Objetivos.....	9
1.2.	Metodología.....	9
1.3.	Etapas	20
1.4.	Material utilizado	22
1.5.	Problemas y dificultades.....	23
2.	Hearing In Noise Test 'HINT'	25
2.1.	Resultados.....	29
2.2.	Líneas futuras	32
3.	Conclusiones	33
4.	Bibliografía	34
A.	ANEXOS	35
a)	Anexo A	35
b)	Anexo B	36
c)	Anexo C	49
d)	Anexo D	49
e)	Anexo E.....	50

1. PRESENTACIÓN Y MOTIVACIÓN

Las personas con deficiencias auditivas tratan de subsanarlas con el empleo de audífonos o implantes cocleares que hoy en día permiten realizar funciones más complejas que la introducción de una simple ganancia, puesto que suelen incorporar pequeños procesadores capaces de realizar complejos cálculos en tiempo real. Esto permite introducir algoritmos que por ejemplo, tratan de reducir el ruido acústico o realzar la señal de voz con vistas a mejorar la inteligibilidad. Sin embargo, el desarrollo de estos algoritmos requiere herramientas de evaluación de sus prestaciones lo más objetivas posible, que a su vez sirvan para calibrar o configurar de forma óptima estos dispositivos.

Según la Real Academia Española, se entiende por ‘Inteligible’, *que se oye clara y distintamente*. Existen muchos factores que afectan directamente a la inteligibilidad. El tiempo de reverberación o la relación señal a ruido, son algunos de ellos cuando consideramos escenarios reales en la comunicación. Este TFM está centrado en cómo afecta la relación señal a ruido en la inteligibilidad de las personas. Para ellos se ha realizado una herramienta que implementa de forma adaptada el test HINT (*Hearing In Noise Test*) [1], donde se presentan distintas frases con distintas relaciones señal a ruido. Esta herramienta tiene el fin de encontrar la relación señal a ruido que permitiría escuchar con un 50 % de inteligibilidad, entendiendo como 50% de inteligibilidad, cuando el 50% de las palabras en una frase son escuchadas y decodificadas correctamente. Otro test que también se utiliza es el SRT (*Speech reception threshold*) [2]. Este procedimiento audiométrico de discurso no invasivo mide el grado de pérdida de la audición para el habla. El umbral de reconocimiento de voz es el nivel de audición más baja a la que el habla apenas puede ser reconocida o comprendida. En esta prueba, un número de palabras *spondaic* se presentan al paciente en diferentes intensidades. Palabras *spondaic* o espondeos, son palabras que contienen dos sílabas que están igualmente acentuadas o enfatizadas cuando se hablan al paciente. El SRT se define como el nivel de audición más baja a la que el paciente repite correctamente 50% de una lista de palabras *spondaic*. Se usan los espondeos, porque son más fáciles de administrar y realizar los test, pero existe el problema de que estas palabras no simulan correctamente una conversación, y no existe naturalidad, por lo que no son la mejor forma de calcular el 50 % de inteligibilidad.

Más del 5% de la población mundial (360 millones de personas), Según datos de la *OMS (Organización Mundial de la Salud)*, padece pérdida de audición discapacitante (328 millones de adultos y 32 millones de niños). Por pérdida de audición discapacitante se entiende una pérdida de audición superior a 40dB en el oído con mejor audición en los adultos, y superior a 30dB en el oído con mejor audición en los niños.

Se dice que alguien sufre pérdida de audición cuando no es capaz de oír tan bien como una persona cuyo sentido del oído es normal, es decir, cuyo umbral de audición en ambos oídos es igual o superior a 25 dB. La pérdida de audición puede ser leve, moderada, grave o profunda. Afecta a uno o ambos oídos y entraña dificultades para oír una conversación o sonidos fuertes.

Las personas ‘duras de oído’ son personas cuya pérdida de audición es entre leve y grave. Por lo general se comunican mediante la palabra y pueden utilizar como ayuda audífonos, subtítulos y dispositivos de ayuda auditiva. Para las personas con una pérdida de audición mayor pueden ser útiles los implantes cocleares.

Las personas ‘sordas’ suelen padecer una pérdida de audición profunda, lo que significa que oyen muy poco o nada. A menudo se comunican mediante el lenguaje de signos.

Las causas de pérdida de audición y sordera se pueden dividir en congénitas y adquiridas.

Las causas congénitas determinan la pérdida de audición en el momento del nacimiento o poco después. La pérdida de audición puede obedecer a factores hereditarios y no hereditarios, o a complicaciones durante el embarazo y el parto, entre ellas: rubéola materna, sífilis u otras infecciones durante el embarazo, bajo peso al nacer, asfíxia del parto, uso inadecuado de medicamentos ototóxicos durante el embarazo o ictericia grave durante el período neonatal, que puede lesionar el nervio auditivo del recién nacido.

Las causas adquiridas pueden provocar la pérdida de audición a cualquier edad.

Algunas enfermedades infecciosas, por ejemplo la meningitis, el sarampión y la parotiditis, pueden ocasionar pérdida de audición, principalmente en la niñez. La infección crónica del oído, que generalmente se manifiesta por supuración ótica, puede causar pérdida de audición. En algunos casos esa infección puede conllevar complicaciones graves, tales como absceso cerebral o meningitis, que pongan en peligro la vida.

La presencia de líquido en el oído (otitis media) puede causar pérdida de audición.

El uso de medicamentos ototóxicos a cualquier edad, incluidos antibióticos y antipalúdicos, puede provocar daños en el oído interno.

Los traumatismos craneoencefálicos o de los oídos pueden causar pérdida de audición.

La exposición al ruido excesivo, por ejemplo el de una maquinaria ruidosa o la música a un volumen muy alto, así como otros ruidos fuertes, como disparos o explosiones, pueden menoscabar la audición.

La pérdida de audición relacionada con el envejecimiento (presbiacusia) se debe a una degeneración de las células sensoriales.

La obstrucción del conducto auditivo producida por cerumen o cuerpos extraños puede causar pérdida de audición a cualquier edad. Ese trastorno suele ser leve y se puede corregir fácilmente.

En los niños, la otitis media crónica es la principal causa de pérdida de audición.

Para poder subsanar estas enfermedades, o profundizar en la investigación sobre ellas, son necesarios estudios previos como el que se presenta en este trabajo final de máster.

La psicofísica es una rama de la psicología que estudia la relación entre los estímulos físicos y las sensaciones o percepciones que generan [3] (Ehrenstein & Ehrenstein, 1999). Dentro de la psicofísica se engloba la psicoacústica, que concentra su estudio en la percepción del sonido. Una de las principales herramientas de trabajo de estas disciplinas son los tests de percepción subjetiva, con los que se recopilan datos de muchos sujetos y se obtienen resultados concretos. Existen diferentes clasificaciones de métodos y tests psicofísicos [4] (Gescheider, 1997) (Nunnally & Bernstein, 1994), pero en psicoacústica está muy aceptada la siguiente clasificación de métodos para cuantificar la impresión sonora [5] (Bech & Zacharov, 2006):

Escala directa: Se basan en pedir al sujeto que realiza la prueba una conversión directa de la sensación sonora en una magnitud sensorial preestablecida en una escala

- Nominal: El sujeto debe decidir si dos estímulos son equivalentes o no.
- Ordinal: El sujeto debe decidir si la magnitud de la sensación es mayor o menor que la de otro estímulo, para un atributo concreto.
- Intervalos: El sujeto debe especificar sobre cada estímulo cuánto le parece mayor o menor con respecto a otro estímulo. Los datos obtenidos darán una lista de ítems ordenados, donde las distancias entre los ítems están especificadas. Los valores absolutos de la escala no son conocidos por los sujetos.
- Ratio: Es similar al método de escala por intervalos con la diferencia de que el valor absoluto de la unidad de medida de la escala sí es conocido.

Escala indirecta: Se mide la sensación mediante la capacidad de los sujetos de discriminar entre distintos estímulos

- Estímulo constante: El sujeto debe indicar cuál de entre dos estímulos da la mayor sensación, incluyendo un estímulo de referencia. Con las respuestas se calcula la probabilidad de que un estímulo sea mayor que otro y a partir de aquí se obtienen umbrales de percepción mínimos.
- Método de los límites: Se buscan los umbrales superior e inferior de percepción a partir de

comparaciones entre una referencia y un estímulo modificado por pasos fijos.

- Método del ajuste: El sujeto debe ajustar una variable del estímulo hasta hacer coincidir su percepción con una referencia.

Existen otras formas de medir la inteligibilidad.

Uno de los modelos de predicción objetivo más ampliamente utilizado es el Índice de Transmisión de la Voz (Speech Transmission Index - STI) [6] desarrollado por Steeneken y Houtgast (1980), [7]. Está pensado para evaluar la deformación de la inteligibilidad de un canal de comunicaciones (o sala). El STI se basa en la generación y análisis de una señal de prueba artificial modulada en amplitud que reemplaza la señal de la voz. Este método supone que la inteligibilidad del habla se mantiene con la acústica de la sala solo si la modulación de la señal artificial se transmite sin alteración desde la fuente hasta la posición del auditor. Como se expresa en la ecuación 1, tanto el tiempo de reverberación T como la relación señal ruido S/R pueden reducir esta modulación. La reducción de modulación en el punto de recepción está asociada con la pérdida de inteligibilidad y su resultado se expresa en un índice cuyo rango va de 0 a 1. La técnica STI está estandarizada por la norma IEC 60268-16 (2003) [8].

$$m(F) = \left[1 + \left(2\pi F \frac{T}{13,8} \right)^2 \right]^{-1/2} \left[1 + 10^{(-S/R)/10} \right]^{-1}$$

Ecuación 1

donde F corresponde a una de las catorce frecuencias de modulación del método.

Un aspecto importante cuando se trata de una prueba de inteligibilidad con voz hablada es el corpus utilizado como material de lectura. Los resultados de las investigaciones realizadas por Hearcom (2005) [9] describen y seleccionan el material disponible para realizar pruebas de inteligibilidad en siete idiomas: danés, holandés, inglés (británico), francés, alemán, polaco y sueco. Los corpus de palabras consideradas en esta publicación están basados en oraciones, palabras monosilábicas (consonante-vocal-consonante - CVC), o tripletas de números. Para el idioma español hay al menos dos corpus de 100 palabras cada uno (Fuchs y Osuna 1965 [10]; Miñana 1969) [11] y fueron estudiados por Sommerhoff y Rosas (2007) [12].

1.1. Objetivos

El objetivo de este trabajo final de máster es definir e implementar una metodología válida para estimar la relación señal a ruido que proporciona un valor del 50% de inteligibilidad para ruidos de distinta naturaleza. La elección o selección de dichos ruidos se realizó cuidadosamente con el fin de que se asemejasen al máximo a los existentes en las situaciones reales que los usuarios se encontrarán en el día a día.

Y siguiendo unas las normas establecidas en el Test Hint [1], se van a manipular los niveles de ruido para calcular el nivel medio de SNR (Signal to Noise Ratio) que necesitan los usuarios para alcanzar un 50% de inteligibilidad. Además nos interesa calcular estos niveles, porque serán el punto de partida a la hora de realizar otros test y estudios.

Se ha desarrollado una aplicación mediante *GUI* de *MATLAB*, que implementa el procedimiento descrito en el test Hint y que se detallará a continuación. Se ha realizado el test en una muestra formada por 30 individuos, con el fin de conseguir una SNR promedio para los distintos ruidos. Partiendo de que tenemos 5 ruidos distintos (ruido de cafetería, ruido de carretera, ruido de estación de tren, ruido aleatorio filtrado con el espectro de la voz masculina, y ruido aleatorio filtrado con el espectro de la voz femenina y que con uno de ellos se aplicaran distintas reverberaciones, cafetería con reverberación media y cafetería con reverberación alta) para no saturar a los encuestados se ha decidido realizar 3 tests por usuario, por lo que al final de las pruebas tendremos 90 encuestas. Se ha realizado un plan de actuación que podemos ver en el ANEXO A, donde aleatoriamente se han asignados distintas combinaciones, con el fin de conseguir un mínimo de 10 test realizados con los ruidos, sin tener que repetir el ruido de cafetería con las distintas reverberaciones a ningún usuario.

1.2. Metodología

Para la realización del test, se han utilizado 5 listas de 10 frases que han sido seleccionadas para que tuvieran sentido, ya que en el Sharvard Corpus al ser una base de datos que tradujo frases del Corpus Ingles, tiene algunas frases que carecen de sentido debido a las diferencias culturales entre la legua inglesa y la española. Obtenidas de la base de datos del *Sharvard Corpus* [13]. Podemos encontrar en el ANEXO D, un total de 50 frases, que serán reproducidas por una voz masculina, y una voz femenina, teniendo un total de 100 frases. Estas frases, han sido ponderadas para ser reproducidas a 64 dB_{SPL}. Para decidir este valor, se ha utilizado un test previo con distintas frases realizado a 5 personas, que consistía en la reproducción de las distintas frases, manipulando el

potenciómetro de la tarjeta de audio, hasta que los encuestados considerasen un nivel de confort parecido al nivel de presión que puede existir en una conversación normal en una cafetería estándar, obteniendo los siguientes resultados: 63 dB_{spl}, 66 dB_{spl}, 62 dB_{spl}, 63 dB_{spl} y 66 dB_{spl}. Promediando los resultados se obtienen 63.8 dB_{spl}, por lo que se ha decidido que el nivel adecuado de confort para presentar las frases del test Hint sea 64 dB_{spl}.



Figura 1 Prueba de audición.

Además se ha analizado el valor *rms* (*Root mean square*) de todas las frases para asegurarnos de que todas ellas tiene niveles similares, obteniendo como resultado promedio 74 dB con una varianza de 4.9 dB.

En la figura 2 podemos ver un par de frases analizadas en detalle, tanto en su evolución temporal, como en el frecuencial. Podemos ver en la imagen una herramienta que ha sido desarrollada como parte del TFM, esta herramienta nos permite visualizar la forma de onda, así como el espectro y el espectrograma. Nos interesa analizar y visualizar estos parámetros para ver donde entra en conflicto con el ruido. Además tiene un selector para poder elegir la frase a analizar. Cuenta con un botón para reproducir la frase, y la frase aparece escrita en pantalla. Más adelante se describe la herramienta de análisis con más detalle.

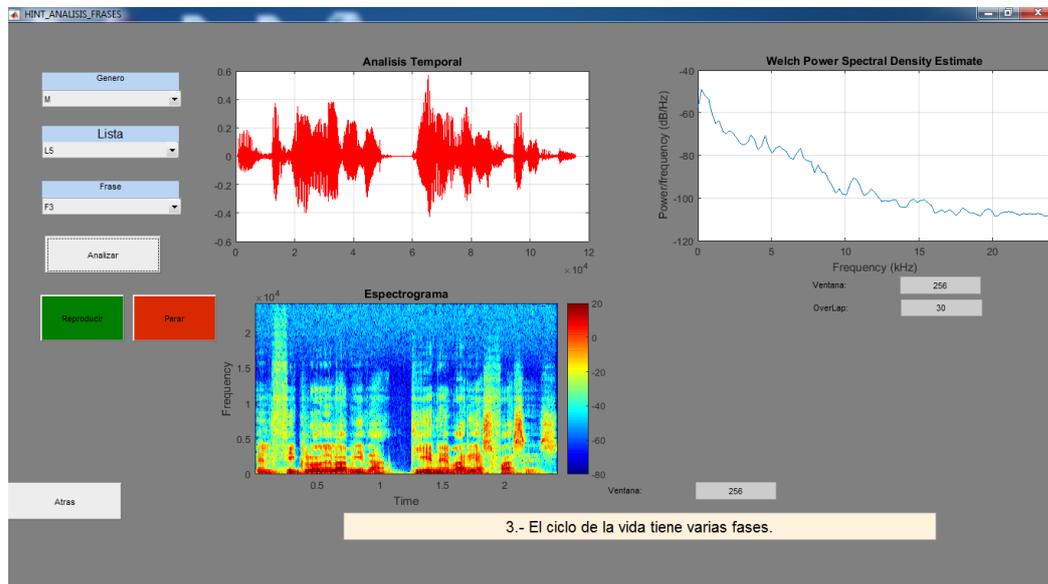
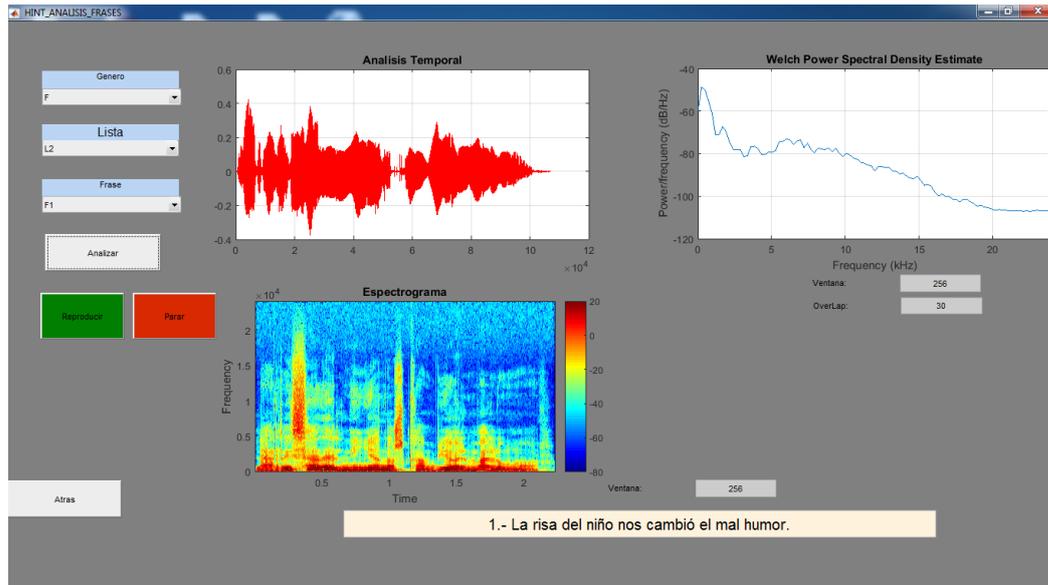


Figura 2: Análisis de un par de frases de la base de datos.

Así mismo se han utilizado 4 ruidos diferentes: ruido de cafetería, ruido de estación de tren, ruido de carretera y ruido blanco, este último, enmascarado con el espectro de la voz tanto masculina como femenina. También para el ruido de cafetería se ha decidido filtrar la voz de dos formas distintas, con una reverberación media y con una reverberación alta, además de la misma sin reverberación alguna. En la figura 4 vemos una herramienta que ha sido programada en este TFM. Esta aplicación es una herramienta de análisis, donde podemos observar tanto la forma de onda, como el espectro y el espectrograma, además de reproducir el ruido en cuestión.

Si observamos los distintos ruidos podemos decir que son señales con altos componentes a baja frecuencia, encontrándose la energía concentrada en el margen frecuencial de 0 a 10 KHz , con una

variación de unos 40 dB y tanto a los 10 segundos como a los 15 segundos se observa unos picos de presión sonora.

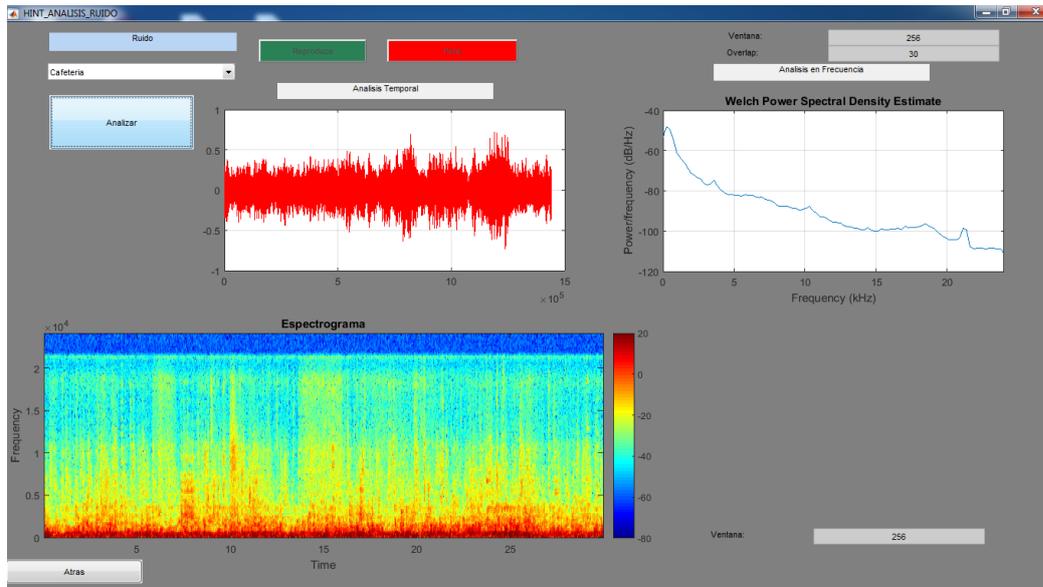


Figura 3: análisis ruido cafetería.

Cuando se observa el ruido de carretera se aprecian distintos matices: las frecuencias predominantes son las bajas (hasta unos 7 KHz se aprecia un nivel elevado de presión sonora). Además si vemos la forma de onda se puede intuir el tráfico de vehículos.

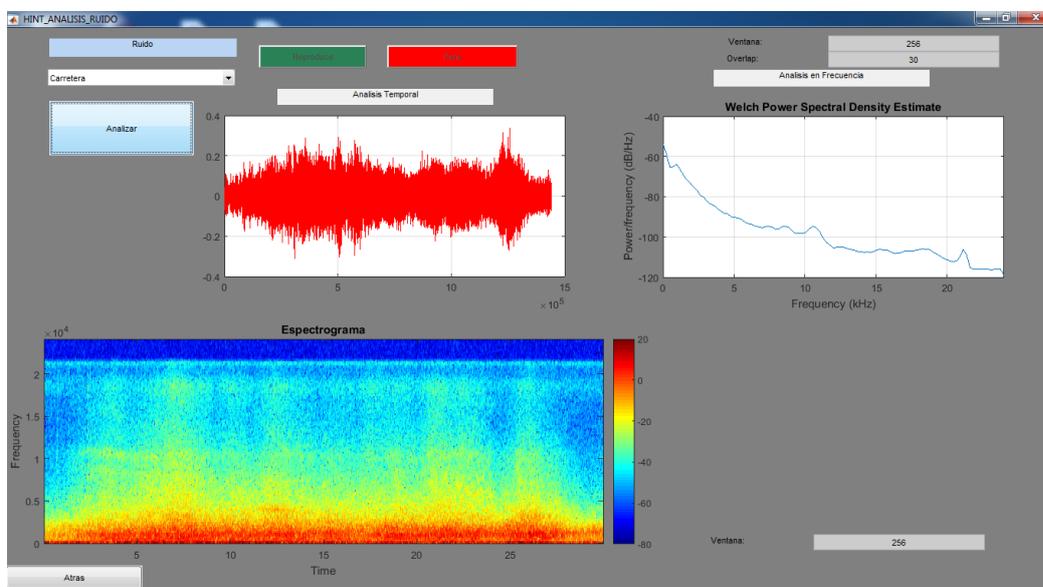


Figura 4: análisis ruido carretera.

El ruido de estación de tren es el que más irregularidades tiene, tanto temporales como frecuenciales, se observan picos en bajas frecuencias, pero además en 10 Khz y en casi los 20 Khz tenemos altos niveles de energía. Temporalmente, a los 10 segundos existe menos nivel, pero antes y después tenemos valores altos de presión sonora.

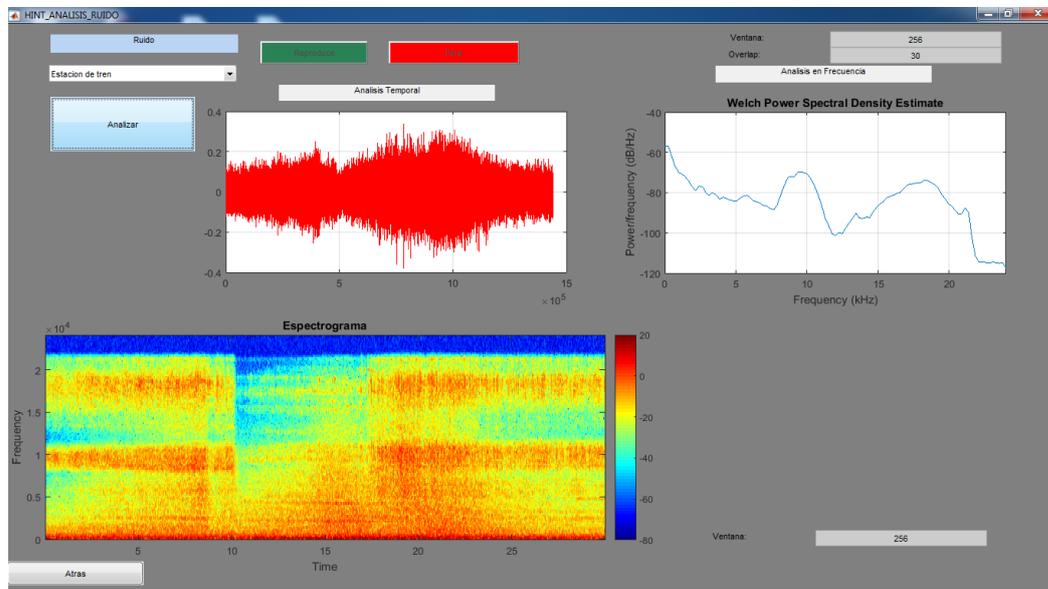


Figura 5: análisis ruido Estación de Tren.

En cuanto a los ruidos aleatorios vemos que son parecidos y que la voz femenina tiene menos potencia a bajas frecuencias.

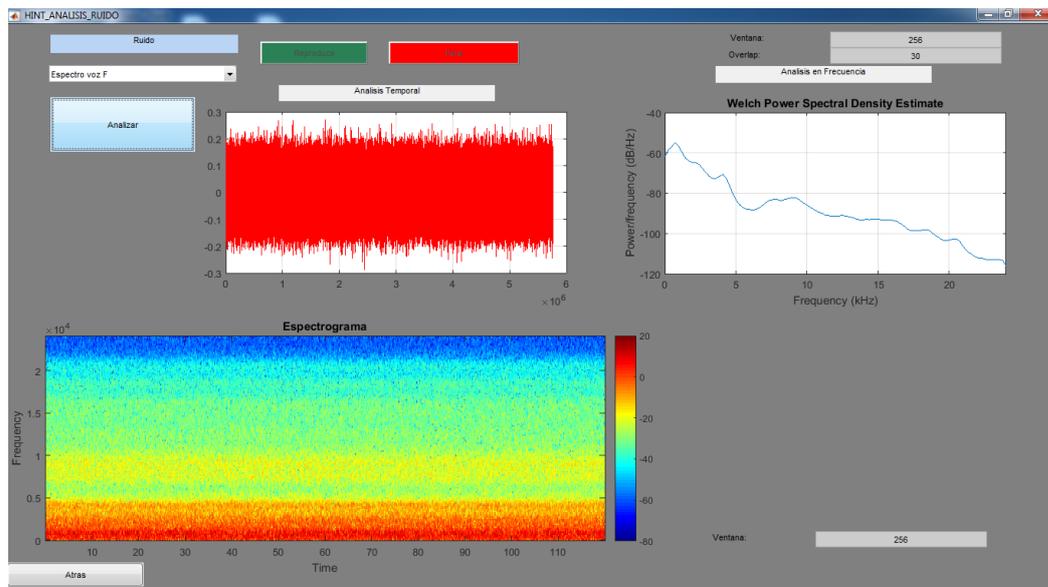


Figura 6: análisis ruido banda ancha con espectro voz femenina.

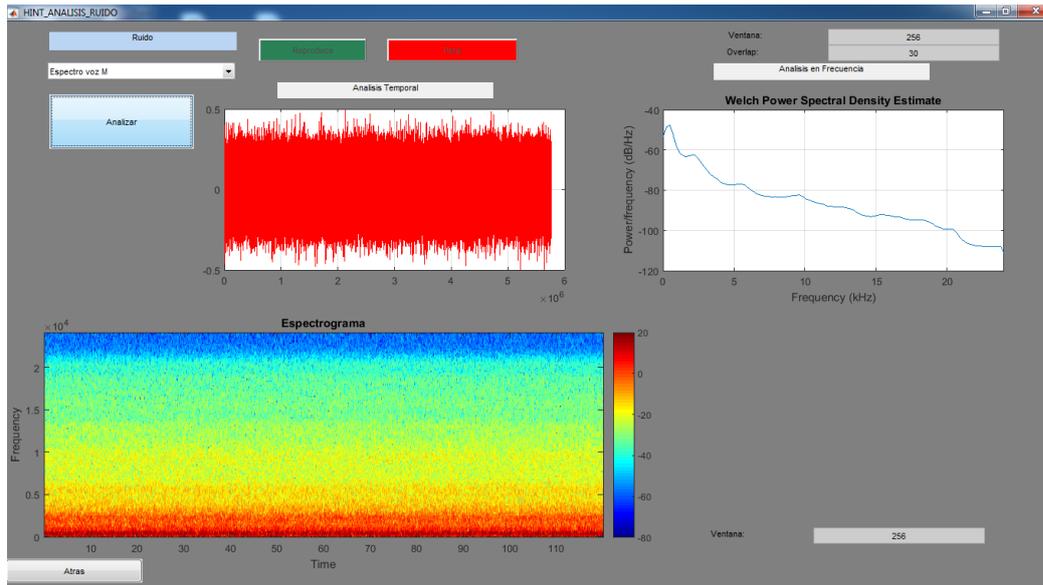


Figura 7 análisis ruido banda ancha con espectro voz masculina.

Para añadir las reverberaciones se ha utilizado el programa de *MATLAB*, *RIR Generator* [14] que sintetiza la respuesta al impulso entre dos puntos de una sala rectangular definida por sus dimensiones y constante de absorción de las paredes. Se simuló una sala de las siguientes dimensiones $x=10\text{m}$ de profundidad, $y=7\text{m}$ de ancho, $z=5\text{m}$ de altura, situando al emisor y al receptor a una distancia de 1.2 m , y fijando para la reverberación media de $\text{TR}_{60}=0.2\text{ s}$ y para la reverberación alta de $\text{TR}_{60}=0.7\text{ s}$.

Con las frases y los ruidos, se ha hecho un croquis de la aplicación que iba a controlar el test.

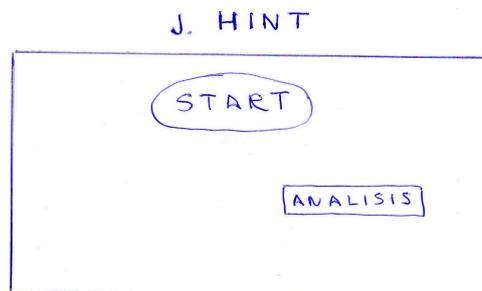


Figura 8 Croquis de la aplicación JHINT.

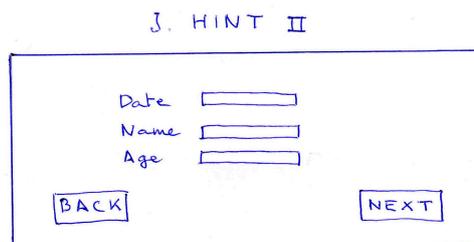


Figura 9 Croquis de la aplicación JHINT, pagina II

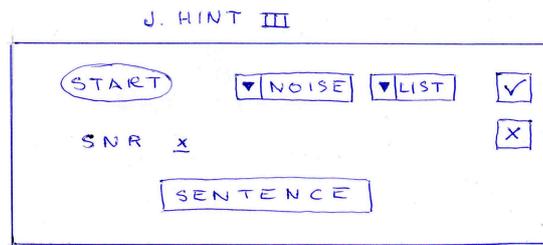


Figura 10 Croquis de la aplicación JHINT, Pagina III

Con el croquis en mente se han decidido algunos puntos clave de la aplicación. La aplicación va a contener varias ventanas enlazadas: en la primera o inicial aparecerá la opción de elegir entre analizar las frases o ruidos, analizar los test realizados, y la posibilidad de empezar el realizar el test.

Se ha añadido un mecanismo de control, para detectar outliers, (encuestados a los que descartaremos para realizar los test) a la hora de empezar los test, para lo que se ha añadido un enlace en la página principal del programa que abre otra aplicación, que permite realizar un simple test de audiometría. Para este test de audiometría se analizaron varias normas, UNE EN ISO 389 parte 1-2-3-4-5-6-7, gracias a esto se ha decidido realizar la audiometría de 125 Hz a 8 KHz. Además se va a añadir una opción para guardar el grafico con el umbral auditivo para el análisis posterior.

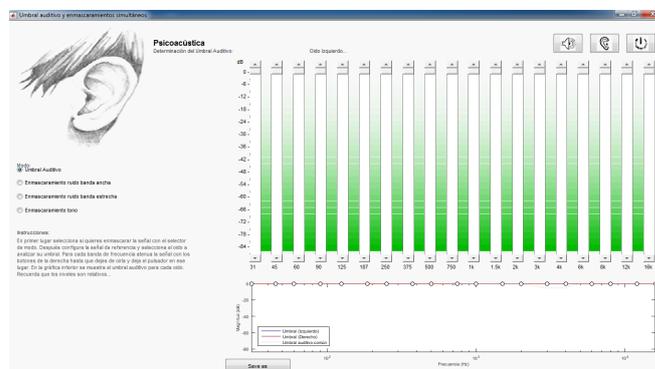
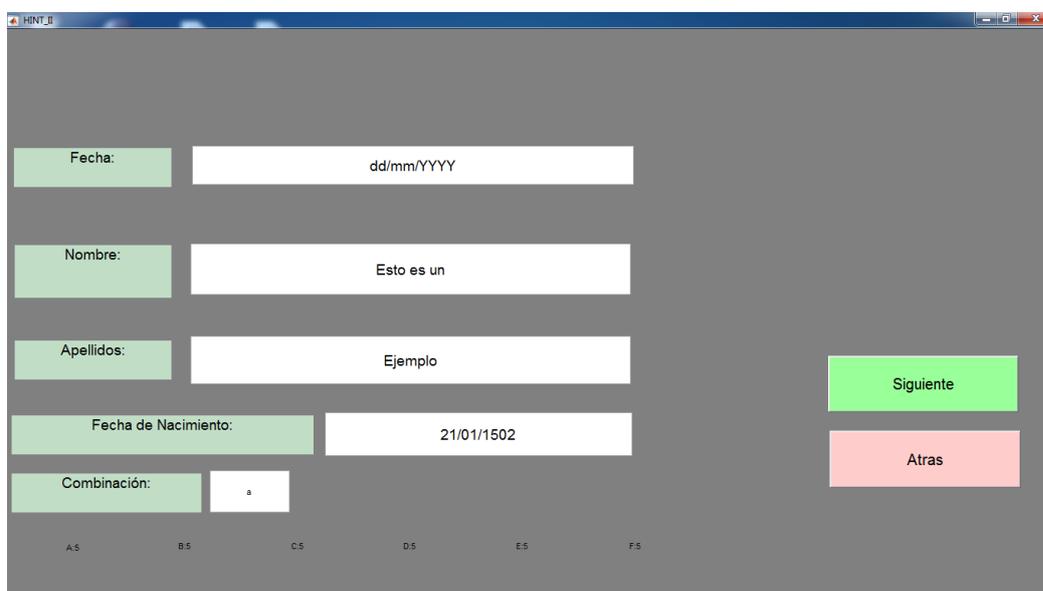


Figura 11 aplicación que realiza el test de audición.

Aunque este será un punto a discutir, ya que realizando varias pruebas, se ha detectado que una leve pérdida de audición no tiene por qué afectar a la inteligibilidad. Para detectar outliers se ha seguido el criterio de la OMS (pérdida superior a 40 dB en el peor oído). Ninguno de los 30 encuestados manifestaron pérdidas superiores a dicho nivel, por lo que no se ha tenido que desestimar a ninguno de ellos.

Una vez realizada la audiometría, tenemos la posibilidad de avanzar a la siguiente fase, donde se recogen los datos de los usuarios a analizar, Nombre, apellidos, fecha de nacimiento y fecha de realización del test. Estos datos nos ayudaran para analizar los resultados por si la edad tuviera algo que ver en la inteligibilidad. Según el plan de actuación, se han elaborado diferentes combinaciones de ruidos y frases, y en esta ventana de la aplicación se ha añadido un contador para saber cuántas combinaciones de cada tipo se han realizado y dar más facilidad a la hora de seguir el plan de actuación.



The screenshot shows a software window titled "HINT II" with a dark grey background. It contains a form with several input fields and buttons. The fields are: "Fecha:" with a placeholder "dd/mm/YYYY"; "Nombre:" with the text "Esto es un"; "Apellidos:" with the text "Ejemplo"; "Fecha de Nacimiento:" with the text "21/01/1502"; and "Combinación:" with a small input field containing the letter "a". To the right of the form are two buttons: a green "Siguiete" button and a pink "Atras" button. At the bottom of the window, there are six small labels: "A.5", "B.5", "C.5", "D.5", "E.5", and "F.5".

Figura 12 segunda página de la aplicación.

El siguiente punto es la prueba de audición donde se elige el tipo de ruido, el género de la voz, y la lista de frases que se reproducirán, atendiendo al plan de actuación.

Con estos puntos elegidos, se empieza el test, según el desarrollo de *Nilsson 1994* [1] sobre el HINT. El nivel de las frases no va a ser modificado, y será el decidido anteriormente. A partir de este la primera frase va a empezar con una SNR de 5dB, ya que como se explica en el artículo de *Nilsson 1994*, es preferible empezar con una SNR favorable para el encuestado y a partir de ésta se va modificando.

El procedimiento detallado para realizar el test HINT, es el siguiente: el usuario va a escuchar la primera frase, y tendrá que repetirla. En el caso de que falle una palabra, se dará el resultado como erróneo, y si acierta todas las palabras se dará el resultado como válido, no penalizando errores en los artículos o en los tiempos verbales. El hecho de que el resultado sea erróneo o valido afectara a la SNR con la que se reproducirá la siguiente frase, repitiendo el procedimiento durante 10 frases, hasta conseguir una SNR final. Los incrementos/decrementos de las SNR son los siguientes: la primera frase, como se ha dicho, a 5 dB, la segunda, tercera y cuarta serán incrementadas o decrementadas (dependiendo de si el usuario entiende correctamente o no dicha frase) en 4dB. Para la quinta, sexta, séptima y octava frase se aplicará un incremento o decremento de 2dB y finalmente para la novena y décima frase la variación de la SNR será de ± 1 dB. Como se puede observar, en cada paso vamos refinando la precisión, para conseguir un resultado final más fiable. Se ha decidido que sean 10 frases por lista, por varios motivos: evitar saturar y acumular de cansancio al encuestado, y porque el hecho de alargarlo con más frases, hace que el usuario, se acostumbre a las condiciones, tanto del ruido como de la voz.

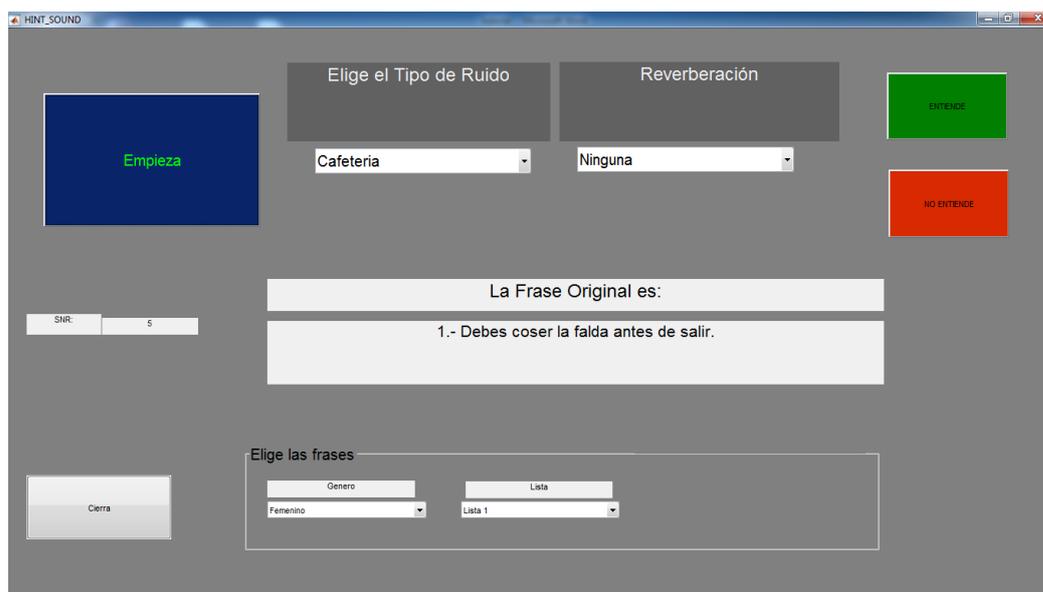


Figura 13 Tercera página de la aplicación.

El procesado realizado para el calculo de SNR es el siguiente: primero se calcula el nivel de presión sonora de la frase, después se calcula el nivel de presión sonora del ruido, mediante la ecuación 2, donde pf es el valor de RMS (*Root Mean Square*) calculada mediante la ecuación 3.

$$Lf = 20 \log_{10} \left(\frac{pf}{2 \times 10^{-5}} \right)$$

Ecuación 2

$$pf = \frac{1}{N} \left(\sum_N f_1^2 + f_2^2 + \dots + f_N^2 \right)$$

Ecuación 3

Calculadas las medias cuadráticas obtenemos el valor SNR a través de la ecuación 4, siendo pf la media cuadrática de la presión sonora de la frase, y pr la media cuadrática de la presión sonora del ruido.

$$SNR = 20 * \log_{10} \left(\frac{pf}{pr} \right)$$

Ecuación 4

El proceso descrito desde la ecuación 2 hasta la 4, implementado en *Matlab* se describe mediante las siguientes instrucciones.

```

Lf=20*log10(pf/2e-5);
Lr=20*log10(pr/2e-5);
pf=(1/length(S))*sum(S.^2);
pr=(1/length(N2))*sum(N2.^2);
SNR=20*log10(pf/pr);

```

A partir de estos niveles fijamos el nivel de ruido para obtener una SNR inicial de 5 dB.

Y dependiendo que frase sea, se deseará fijar una u otra SNR, como se describe en el código de *Matlab*. Utilizando un condicional *if* calculamos la SNR deseada.

```

if frase<=5
    SNRDESEO=(SNRD-4)/2;
else if frase>5&&frase<=9
    SNRDESEO=(SNRD-2)/2;
    else if frase>=10
        SNRDESEO=(SNRD-1)/2;
    end
end
end
end

```

El siguiente punto es calcular el factor que multiplicara la señal de ruido. En el código de *Matlab* que hay a continuación vemos el proceso. Una vez obtenido el factor, multiplicamos por el ruido, aumentando o disminuyendo el valor de presión sonora según el factor calculado.

```
PRS=pf/(10^(SNRDESEO/10));  
factor=PRS/pr;  
factor=sqrt(factor);  
N2=N2*factor;
```

Paralelamente se ha trabajado en una herramienta de análisis dentro de la misma aplicación donde podemos estudiar distintos parámetros: por un lado la evolución temporal y frecuencial de los ruidos y las frases, , y por otro lado, los resultados de los test realizados. En la figura 14 podemos ver un boceto de la idea para esta herramienta. Más adelante se describirá con profundidad.

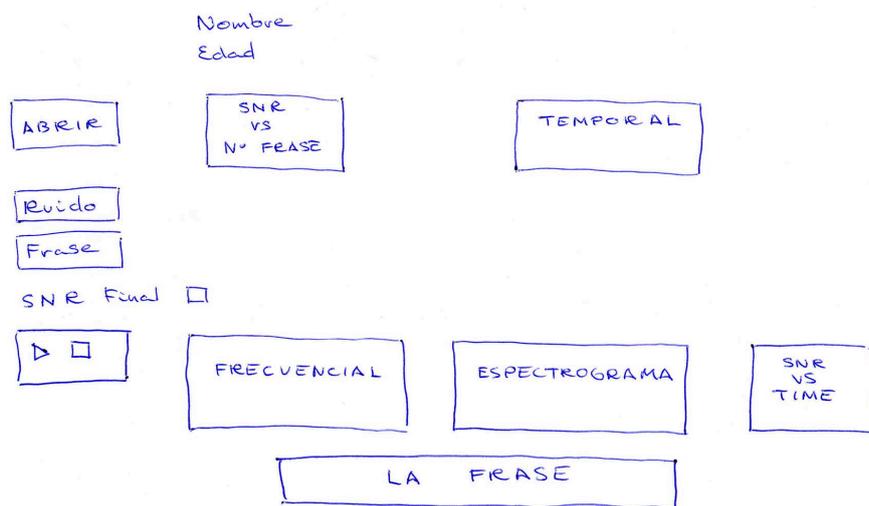


Figura 14 Croquis de la aplicación
JHINT, Analizador

Finalmente, como vemos en la figura 15, las prestaciones de la herramienta de análisis son las siguientes: abrimos el test que queremos analizar, seleccionamos el ruido y la frase que queremos

ver, aparece la SNR final, La SNR de la frase, La SNR vs N° de frase, un análisis temporal, un análisis espectral, el espectrograma, la evolución temporal de la SNR, y además se añadió, la evolución de la potencia con el tiempo de la frase y la evolución de la potencia con el tiempo del ruido, incluyendo la posibilidad de visualizar el audiograma realizado al usuario al que se está analizando. La herramienta permite modificar varios parámetros para el análisis como el número de puntos de la fft (*Fast Fourier Transform*).

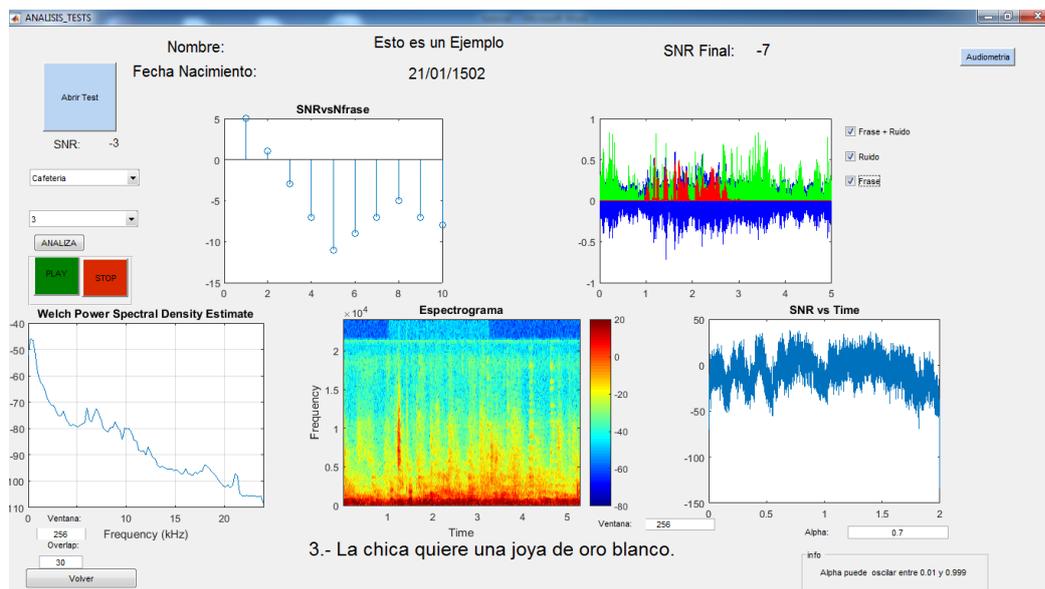


Figura 15 Aplicación JHINT, Analizador

1.3. Etapas

Las etapas del proyecto se pueden diferenciar en 4 grandes fases: investigación, diseño de la aplicación, realización de los test y análisis de los resultados junto con la redacción de esta memoria.

El primer punto a abordar fue la búsqueda de información e investigación sobre el tema en cuestión, la lectura de artículos, como [1], donde se explica el desarrollo del test HINT, o el *Sharvard Corpus 2014* [13], donde se realiza un estudio con frases en lengua castellana desde la Universidad del País Vasco y la University of Western Sydney, que complementan a [15].

La segunda fase del trabajo, fue el diseño de la aplicación y la realización de la misma en *MATLAB*. Esta fase del proyecto es la que ha demandado más tiempo para su realización ya que el

diseño es la parte más importante para conseguir que todo sea más fluido, previendo los problemas que puedan surgir, para poder abordarlos desde un principio. Al mismo tiempo, la realización de la aplicación y su programación fue una tarea intensa del proyecto por la cantidad de tiempo invertido en la misma. Para ello se investigó sobre la aplicación *GUI* de *MATLAB* para que la aplicación tuviera un entorno amigable, y además funcionara de forma eficiente en todos sus aspectos. En el ANEXO C se puede ver el código realizado para la aplicación.

La tercera parte del proyecto es la realización de los test a diferentes individuos. Se realizó un plan de actuación para que encuestando al menor número de usuarios, se obtuvieran suficientes muestras para poder hacer un análisis con todos los ruidos de la forma más eficiente. Se estimó que lo óptimo era encuestar a 30 individuos con el fin de obtener un mínimo de 10 test por ruido. Cada persona realizó 3 tests para evitar la fatiga de estas. Podemos ver el plan de actuación en el ANEXO A.

La última fase del proyecto, además de la redacción del mismo, fue el análisis de los resultados. Utilizando el entorno de *STATGRAPHICS* se ha hecho un análisis exhaustivo de los test realizados. Podemos ver en el apartado de resultados los análisis de interés, y además en el ANEXO D está el análisis más ampliado con todos los factores analizados.

Para la realización de este proyecto se han utilizado aptitudes y conocimientos adquiridos en el Master de ingeniería Acústica. Como la programación en *MATLAB*, el estudio estadístico con *STATGRAPHICS*, la lectura de textos científicos y normas UNE en ISO, la utilización de instrumental científico como sonómetro o cabeza acústica... entre otros.

1.4. Material utilizado

La aplicación ha sido compilada en un archivo ejecutable, por lo que puede ser ejecutada en cualquier ordenador, con sistema operativo *Windows*. Además se ha añadido un enlace de descarga de un complemento para ordenadores que no tengan *matlab* instalado, por lo que no es necesario tenerlo instalado.

Para la realización de este proyecto se ha necesitado la utilización del siguiente instrumental científico y recursos:

Equipo personal Samsung con procesador Intel Core i3-2328M CPU @ 2.2GHz, con 8GB de Memoria RAM, sistema operativo a 64 bits con entorno *Windows 7*. Además, se usó una tarjeta de audio M-Audio Fast Track y unos auriculares Sennheiser hd 250. Con este equipo tenemos todo lo necesario para realizar los test. También se ha utilizado un sonómetro de Brüel & Kjaer type 2250, calibrador Type 4231, y una cabeza acústica de Brüel & Kjaer type 4100 D, necesario tanto para calcular el ruido de fondo de la sala donde se realizaron los test como para calibrar el equipo.

También se utilizó una sala donde el ruido de fondo es menor a 30 dB_{spl}.



Figura 16 material utilizado.



Figura 17 Sonómetro y calibrador.



Figura 18 Tarjeta de audio.

1.5. Problemas y dificultades.

La presencia de problemas durante la realización de un proyecto ayuda a perfeccionar el trabajo y además a aprender. Durante el transcurso de este proyecto fueron apareciendo varios problemas los cuales van a ser comentados detalladamente.

Una vez finalizada la aplicación existen varios detalles a tratar, como la calibración de la instrumentación, para asegurarnos que las frases son emitidas a 64 dB_{spl}. Además para asegurar que la audiometría se realiza a partir de unos umbrales según la normativa, es necesaria la calibración de los dispositivos. Para esto se consultó la norma UNE_EN_60942 2005, la UNE_EN_61094-2 2011, la UNE_EN_ISO_389 parte 1,2,3,4,5,6,7,8,9, la UNE_EN_ISO_8253-1 2011 y la norma UNE_EN_60318-1 2010.

Después de consultar las normas citadas anteriormente, mediante un sonómetro y la cabeza acústica se ha calibrado el equipo.



Figura 19 calibración
de auriculares.



Figura 20 sonómetro auriculares

Otro problema que tuvo que ser abordado fue la adecuación del margen dinámico de las señales emitidas por la aplicación desarrollada en *Matlab*, para evitar saturaciones en los niveles de reproducción. Para ello, fue necesario encontrar el peor caso y normalizar respecto al mismo. Una vez realizada la normalización se ha asegurado que todas las frases fueran emitidas al valor de 64 dB_{spl} .



Figura 21 Nivel del potenciómetro de la tarjeta de audio.

A la hora de realizar los test, es necesario asegurarnos de que el ruido de fondo no supera los 30 dB_{spl} , así que se ha calculado el ruido de fondo empleando un sonómetro para asegurarnos de que se cumplen las condiciones preestablecidas.



Figura 22 Calculo del nivel de ruido de fondo.

Otro problema es la naturaleza determinista de las señales de ruido usadas, ya que, por ejemplo, a la hora de reproducir algunos ruidos como el de carretera, siempre y al final del mismo se escucha un coche que perjudica el final de la frase, así pues se ha decidido reproducir el ruido aleatoriamente a partir de distintos puntos.

Finalmente, solucionados estos problemas, se ha realizado el test a los individuos y se ha procedido al análisis de los resultados.

2. HEARING IN NOISE TEST 'HINT'

La Prueba de Audición en Ruido (HINT) es una prueba de recepción del lenguaje estandarizado que mide el reconocimiento de frases con ruido de fondo. El HINT se compone de 250 frases, que se clasifican en 25 listas. Las oraciones para el HINT son una adaptación de 336 frases [16].

El HINT es una tarea de reconocimiento de voz adaptativa. En otras palabras, el nivel de cada ruido se ajusta basándose en la respuesta del oyente. El nivel de presentación del ruido aumenta después de cada respuesta correcta, lo que eleva el nivel de dificultad de la siguiente frase en la lista. Por el contrario, el nivel de presentación se disminuye después de cada respuesta incorrecta, lo que reduce la dificultad para la siguiente frase. El nivel de reproducción de las frases se mantiene constante, por lo que adaptando el nivel de presentación del ruido varía la relación de señal a ruido (SNR).

Para determinar si una respuesta es correcta o incorrecta, al encuestado se le expone una frase, y este debe reproducirla, repitiendo exactamente lo que ha escuchado. El HINT especifica que para determinar que una frase es correcta, el encuestado debe acertar el 100% de las palabras de esta, y que se puede fallar en la reproducción de los artículos, y además en la reproducción de los tiempos verbales.

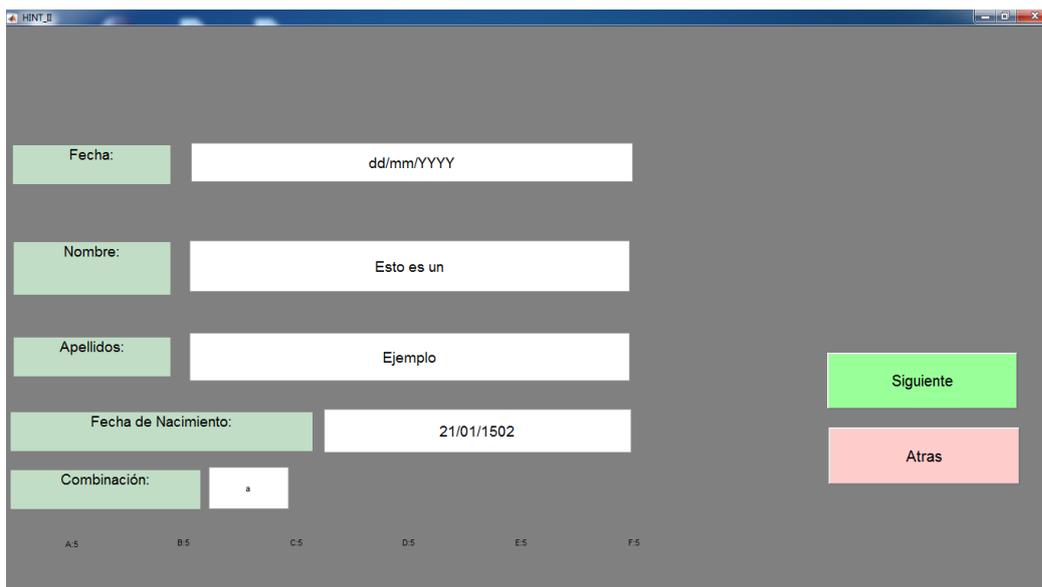
Hay que tener en cuenta que las frases HINT También se pueden administrar sin ruido para evaluar el reconocimiento de las frases en silencio. En este caso, se obtiene un umbral para el reconocimiento de frases. Si se realiza la prueba con el ruido, la adaptación de la SNR permite la estimación de un umbral de SNR para el reconocimiento del habla en presencia de ruido. Cuando el valor de la relación SNR disminuye, las condiciones de escucha se vuelven más difíciles.

Por ejemplo, si un oyente con problemas de audición no es capaz de entender el habla con una SNR de -3 dB usando un audífono omnidireccional pero puede ser capaz de entender el mensaje con una SNR de -6 dB cuando el ajuste del micrófono se ha cambiado a direccional, el cambio en el

micrófono permite al oyente entender el habla en las condiciones más adversas, mejorando así su capacidad de reconocimiento de voz con el ruido.

El Hint utiliza ruido aleatorio para la realización de las pruebas, pero en este TFM se han utilizado ruidos cotidianos para tener unos resultados más veraces. En este caso es necesario implementar el Hint para sacar el resultado concreto de la SNR para los distintos ruidos, ya que con estas SNR luego se podrán realizar otros test como los basados en el artículo de Lena L. N. Wong (2012) [17].

La aplicación Hint que se ha realizado en este TFM, cuenta con distintas particularidades. Esta realizada en un entorno amigable, por lo que cualquier persona no familiarizada con la programación puede utilizarla. El programa cuenta con un diseño particularizado exclusivamente para esta aplicación. Si observamos la página 2 del programa, figura 23, advertiremos que se nos piden ciertos datos como la fecha actual y la de nacimiento del encuestado, para poder calcular las edades, y tener un control de cuando se han realizado los tests, o el nombre y apellidos del encuestado, para almacenarlo en una base de datos. Además el campo combinación se ha creado para poder llevar un recuento del plan de actuación.



The screenshot shows a software window titled "HINT_II" with a dark gray background. The form contains the following elements:

- Fecha:** A text input field with the placeholder "dd/mm/YYYY".
- Nombre:** A text input field with the placeholder "Esto es un".
- Apellidos:** A text input field with the placeholder "Ejemplo".
- Fecha de Nacimiento:** A text input field with the placeholder "21/01/1502".
- Combinación:** A text input field containing the character "a".
- Navigation:** A green button labeled "Siguiente" and a pink button labeled "Atras".
- Footer:** A row of small text labels: "A5", "B5", "C5", "D5", "E5", "F5".

Figura 23 Pagina II del JHINT.

Todos estos campos son almacenados en una variable de *matlab* tipo *cell* y guardadas en un archivo tipo *.mat llamado Test.mat donde irán almacenados todos los test para así poder luego acceder a ellos para analizarlos. Este archivo *.mat contiene los siguientes campos: Fecha, Nombre, Nacimiento, Doc(Documento), y Comb(combinación).

'Fecha'	'Nombre'	'Nacimiento'	'Doc'	'Comb'
'10/07/2015'	'Jorge Enrique Otero Vega'	'14/03/1986'	'Jorge Enrique Otero Vega'	'a'
'13/07/2015'	'Juan Estreder Campos'	'03/06/1991'	'Juan Estreder Campos'	'a'
'13/07/2015'	'Jorge Lorente Giner'	'22/03/1985'	'Jorge Lorente Giner'	'a'

Tabla 1 Test.mat

El archivo Doc nos sirve para luego enlazar con los resultados del encuestado en cuestión. Cuando pasamos al siguiente punto del programa. Donde se realizan las pruebas, el programa está diseñado para que guarde los datos de los resultados en un archivo *.mat, donde podemos observar los siguientes campos, Doc, Ruido, Lista,SNR1..SNR10,SNRfinal. De este modo después en el analizador solo hay que buscar el nombre a analizar y aparecen los resultados, con la herramienta que se ha desarrollado explícitamente para ello.

'Doc'	'Ruido'	'Lista'	'SNR 1'	'SNR 2'	...	'SNR 9'	'SNR 10'	'SNR Final'
'Sandra Marin Martinez'	'Cafeteria'	'FLista3'	'5'	'1'	...	'-7'	'-6'	'-5'
'Sandra Marin Martinez'	'Estación de tren'	'MLista4'	'5'	'1'	...	'-3'	'-4'	'-3'
'Sandra Marin Martinez'	'Ruido aleatorio espectro F'	'FLista5'	'5'	'1'	...	'-3'	'-2'	'-1'

Tabla 2 Tests.mat

En la figura 24 podemos observar la herramienta que se ha diseñado para el análisis de los diferentes test.

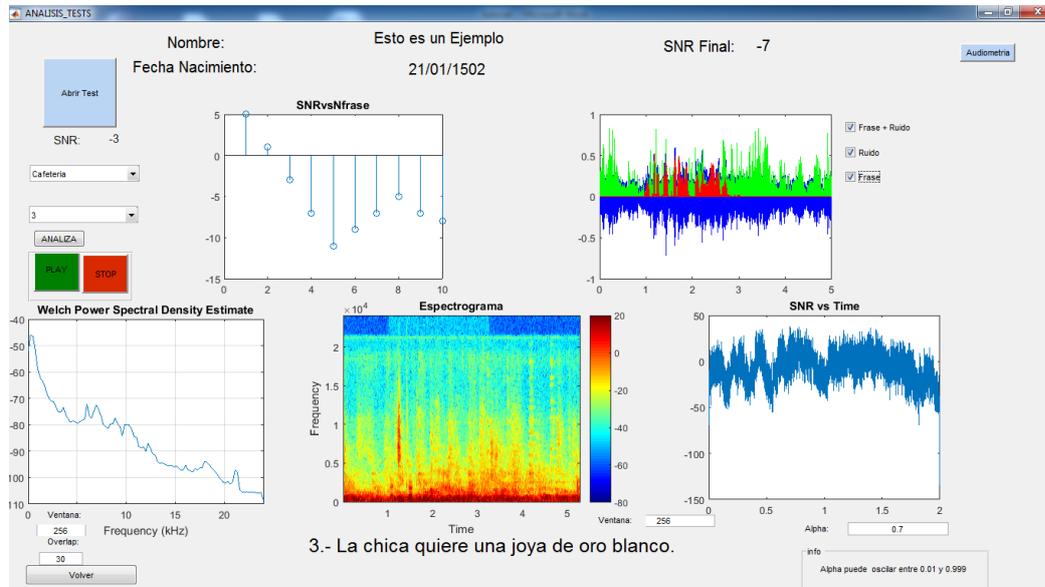


Figura 24 analizador del JHINT.

Si observamos la herramienta podemos advertir que nos ofrece muchas prestaciones. Para empezar, pulsando el botón de abrir test, podemos seleccionar el archivo *.mat con el nombre del usuario a analizar. Una vez abierto se carga el nombre y la fecha de nacimiento del encuestado. Pulsando en el botón de audiometría, podemos abrir una imagen en *.fig donde se observa el umbral de audición realizado al encuestado. La herramienta dispone de dos botones desplegables: en el primero se puede elegir entre los ruidos con los que se le ha encuestado al usuario. El segundo, ofrece las 10 frases con las que cuenta cada test. Una vez abierto un test, seleccionada una frase y pulsando el botón de analiza, podemos ver lo siguiente: en la primera figura aparece una gráfica con las 10 frases y sus SNR, donde arriba a la derecha se puede ver la SNR final para ese ruido, y arriba de los botones desplegables, aparece la SNR para la frase seleccionada. Además la aplicación cuenta con distintas gráficas, que permiten visualizar la forma de onda, pero si se pulsan los botones *ruido o*frase, aparecen sobre la gráfica la evolución de la potencia frente al tiempo de estas. Por ultimo las gráficas de la parte inferior ofrecen el espectro de la señal, un espectrograma, y la más interesante una gráfica con la evolución temporal de la SNR, por lo que podemos ver cuánto tiempo o en qué momento la SNR era demasiado alta, o por lo contrario baja, ya que muchas veces solo afecta a una palabra, o puede ser que a ninguna.

2.1. Resultados

La realización de estos test pretende obtener unos resultados específicos del nivel de SNR que proporcionaría el 50% de inteligibilidad para los distintos ruidos preseleccionados. Se espera, que los resultados se ajusten en función de la naturaleza de cada ruido.

En la tabla 3 podemos observar los resultados obtenidos

Resumen Estadístico

Resumen Estadístico para Snr final

<i>Ruidos</i>	<i>Recuento</i>	<i>Promedio</i>	<i>Mediana</i>	<i>Moda</i>	<i>Varianza</i>	<i>Desviación Estándar</i>
'Cafetería Ra'	10	-2,8	-3,0	-3,0	4,8444	2,201
'Cafetería Rm'	10	-2,8	-3,0	-3,0	27,956	5,2873
'Cafetería'	10	-5,4	-5,0	-5,0	5,1556	2,2706
'Carretera'	20	-7,4	-7,0	-5,0	12,042	3,4702
'Estación de tren'	20	-6,9	-7,0	-9,0	9,0421	3,007
'Ruido aleatorio espectro F'	13	-1,6154	-1,0	-1,0	7,5897	2,7549
'Ruido aleatorio espectro M'	7	-3,8571	-5,0	-5,0	2,4762	1,5736
Total	90	-4,9333	-5,0	-3,0	14,512	3,8095

Tabla 3 Resumen estadístico

Para el ruido de cafetería con una reverberación alta se han realizado 10 test, obteniendo una SNR promedio de -2,8 dB, el valor de la mediana es de -3 dB y la Moda es de -3dB. Puesto que existe una desviación estándar de sólo de 2,20 (considerando los 10 encuestados) consideramos que este resultado es un indicativo de que la SNR promedio es fiable.

Si observamos los resultados del ruido de cafetería con reverberación media, obtenemos unos resultados similares: una SNR de -2,8 dB y una moda de -3 dB, sin embargo la desviación típica es de 5,28. Esto es debido a que un sujeto dista mucho de los demás encuestados, por lo que se ha decidido descartar el resultado del sujeto en cuestión y realizar el mismo test a otro usuario. con lo

que obtenemos el siguiente resultado: una SNR promedio de -4,4dB y una desviación estándar de 2,11.

Resumen Estadístico para Snr final

Ruidos	Recuento	Promedio	Mediana	Moda	Varianza	Desviación Estándar
'Cafetería Ra'	10	-2,8	-3,0	-3,0	4,8444	2,201
'Cafetería Rm'	10	-4,4	-4,0	-3,0	4,4889	2,1187
'Cafetería'	10	-5,4	-5,0	-5,0	5,1556	2,2706
'Carretera'	20	-7,4	-7,0	-5,0	12,042	3,4702
'Estación de tren'	20	-6,9	-7,0	-9,0	9,0421	3,007
'Ruido aleatorio espectro F'	13	-1,6154	-1,0	-1,0	7,5897	2,7549
'Ruido aleatorio espectro M'	7	-3,8571	-5,0	-5,0	2,4762	1,5736
Total	90	-5,1111	-5,0	-3,0	11,628	3,41

Tabla 4 Resumen estadístico

Para el ruido de cafetería sin reverberación, con 10 encuestados tenemos los siguientes resultados: promedio:-5,4dB, moda de:-5 dB y desviación estándar:2,27.

Con el ruido de carretera, se han encuestado a 20 usuarios, y la SNR promedio es de -7,4 dB, con una moda de -5 dB y una desviación típica de 3,47.

Observando los resultados para el ruido de estación de tren, con 20 encuestados, existe una SNR promedio de -6,9, la moda es de -9 y la desviación estándar de 3,00.

Para el Ruido de espectro ensanchado filtrado con el espectro de la voz de la mujer, la SNR promedio es de -1,61dB, con una moda de -1dB, y una desviación estándar de 2,75.

Y para el Ruido de espectro ensanchado filtrado con el espectro de la voz del hombre, la SNR promedio es de -3,85 y una desviación estándar de 1,57.

Como era de esperar los ruidos que fueron procesados con cierto nivel de reverberación, son los que necesitan una SNR mayor, además del de ruido de espectro ensanchado Estos son los ruidos que los usuarios encuestados han encontrado mayor dificultad a la hora de entender el mensaje.

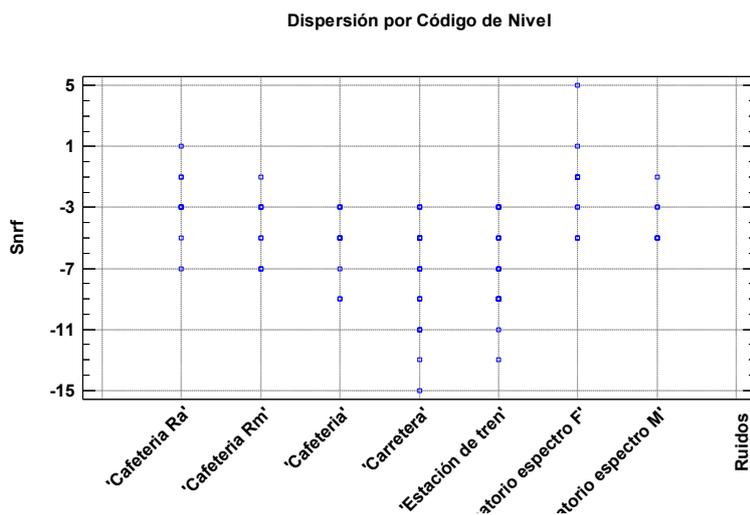


Figura 25 Dispersión por Ruido.

En la figura 25 podemos observar la dispersión para cada ruido. Se observa que para el ruido de carretera tenemos una mayor dispersión, ya que la presencia de este ruido ha originado en los encuestados más dudas.

Si observamos la figura 26, se puede observar que para el ruido de carretera existe una desviación mayor, aunque el resultado final es estable, sin embargo para el ruido de cafetería con reverberación alta existe una muestra que está muy alejada del promedio, al igual que pasa con el ruido aleatorio con espectro femenino, donde también existe una muestra con Snr Final que es igual a 5, estando muy desplazada del promedio.

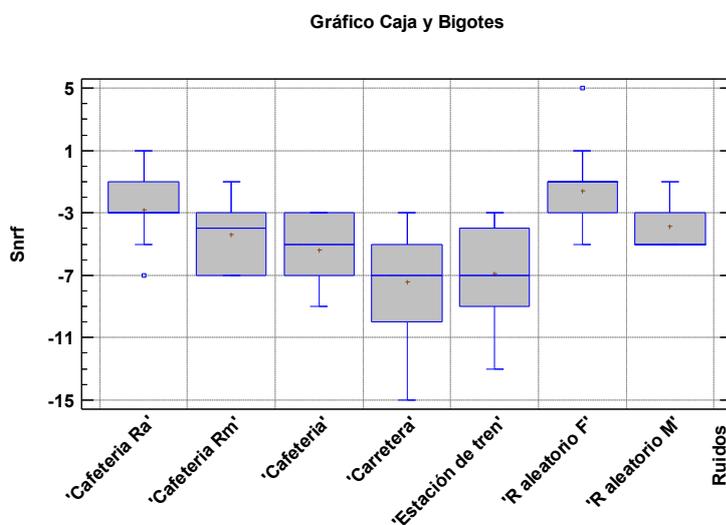


Figura 26 Grafico de Cajas y bigotes.

2.2. Líneas futuras

El trabajo realizado y comentado en este documento nos permite calcular el nivel SNR para el cual obtenemos un 50% de inteligibilidad. Aunque ha sido aplicado a los distintos ruidos descritos a lo largo del trabajo, una posible línea futura sería utilizarlo con otros ruidos cotidianos que fueran de interés.

Por otra parte esta caracterización de la inteligibilidad afectada por ruidos de diferente naturaleza, nos serviría para mejorar las prestaciones de los audífonos, facilitando el día a día a las personas con deficiencias auditivas.

Además, usando esta herramienta podremos conseguir configurar adecuadamente los audífonos, ya que puede emplearse en usuarios de estos dispositivos. Por ejemplo, si le realizamos el test a una persona que utiliza un audífono, y calculamos que para un ruido determinado necesita una determinada SNR, después podemos aplicarle un filtro al audífono y volver a realizar el test, para conseguir reducir la SNR y aumentar las cotas de inteligibilidad. Por tanto, a través del análisis de las diferentes configuraciones de los audífonos realizados mediante esta aplicación se podrían evaluar las mejoras en el procesado de dichos audífonos.

Entre otras muchas líneas futuras, destacamos por ejemplo que a partir de los resultados obtenidos podemos tratar de emular el test que realiza *Lena L.N. Wong 2012*, [17] donde se caracteriza la inteligibilidad para distintos ruidos pero de forma más rigurosa, buscando otras cotas de inteligibilidad como el 20%, 40% u 80% aunque particularizado con la lengua China cantonesa. En nuestro caso partiendo de la SNR obtenida para el 50% de inteligibilidad se podrían extrapolar estos resultados con el fin de encontrar un patrón en ellos pero, utilizando el corpus español.

En la misma línea, tratando de mejorar la inteligibilidad se puede ajustar el test que se ha diseñado pero usando otras configuraciones de forma que nos permita analizar la inteligibilidad de los encuestados, considerando palabra a palabra y no con frases completas. Para esto, basándonos en [17] se deberá realizar una lista con frases perfectamente balanceadas, otras donde predominen las palabras sordas y otras donde predominen las palabras sonoras. Si se analizan todos los parámetros se podrá caracterizar con mayor precisión la inteligibilidad, y así intentar buscar en función de la naturaleza del ruido, una configuración genérica (procesado a realizar en los audífonos) que optimice la capacidad de entendimiento.

Por ultimo, mencionar que se podría mejorar y actualizar la herramienta realizada en este proyecto, para permitir, por ejemplo, un análisis en tiempo real, con objeto de poder detectar errores durante la realización de las encuestas.

3. CONCLUSIONES

En este trabajo final de master se ha estudiado procedimientos para calcular la inteligibilidad mediante percepción del habla, lo que nos ha conducido a realizar una herramienta que permite calcular el nivel de SNR para conseguir el 50 % de inteligibilidad. Esta herramienta permite evaluar la inteligibilidad no utilizando ruido blanco como señal interferente tal y como se describe en [1], sino empleando ruidos procedentes de la realidad cotidiana. Como se preveía, el nivel de SNR obtenido varía en función de la naturaleza del ruido, como era de esperar (por ejemplo, necesitamos mayor SNR para obtener la misma inteligibilidad, conforme los ruidos son mas aleatorios o para las frases con mayor reverberación).

Además si comparamos los resultados con los obtenidos por *Nilsson 1994 [1]* (donde se usaron ruido blanco y balbuceo de gente), comprobamos que para el caso del ruido aleatorio obtenemos unos resultados similares, por lo que queda contrastada la validez de la metodología utilizada.

Podemos afirmar que los valores de SNR obtenidos para los distintos ruidos considerados son fiables, puesto que la dispersión obtenida tras realizar el estudio estadístico proporciona una varianza dentro de unos limites tolerables.

Destacar que la aplicación del test Hint realizada en este TFM ha sido diseñada desde el principio considerando las directrices de *Nilsson 1994 [1]*, lo que ha conllevado a planificar y tomar decisiones (relacionadas con la programación, temporización,...) en función de dichos objetivos iniciales y de los recursos disponibles.

La realización de este TFM me ha ayudado a reforzar los conocimientos adquiridos durante el Master de ingeniería Acústica, ya que se ha trabajado en programación en *Matlab* para realizar la aplicación, en tratamiento de la señal, para el análisis de las frases, ruidos y test. También en análisis estadístico a la hora de analizar los resultados, además de consultas bibliográficas, material técnico, como el sonómetro o la cabeza acústica, la lectura y comprensión de normas UNE EN ISO, y otras disciplinas transversales.

Ha de valorarse que este trabajo ha sido aceptado para una comunicación en el congreso de TecniAcústica que será celebrado en Valencia, en octubre del 2015.

4. BIBLIOGRAFÍA

- [1] Nilsson, M., Soli, S.D., & Sullivan, J.A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and noise. *Journal of the Acoustical Society of America*, 95(2), 1085-1099.
- [2] ANSI. (1997) *S3.5 American National Standard Methods for Calculation of the Speech Intelligibility Index*. (Acoustical Society of America, New York)
- [3] Ehrenstein, W; & Ehrenstein, A. (1999). Psychophysical Methods. In U. Windhorst, & H. Johansson, *Modern Techniques in Neuroscience Research* (pp. 1211-1241). Berlin: Springer-Verlag Berlin Heidelberg.
- [4] Gescheider, G. (1997). *Psychophysics: The fundamentals*. Mahwah, Nueva Jersey: Lawrence Erlbaum Associates Inc.
- [5] Bech, S; & Zacharov, N. (2006). *Perceptual Audio Evaluation-Theory, Method and Application*. Sussex, Inglaterra: John Wiley & Sons Ltd.
- [6] Steeneken, H.J.M y T. Houtgast. 1980. "A physical method for measuring speech transmission quality". *J. Acoust. Soc. Amer* 67 N° 1.
- [7] Steeneken, H.J.M. y T. Houtgast. 2002. "Validation of the STIr method with the revised model". *Speech Communication* 38: 413-425.
- [8] International Standard Organization. "Ergonomics - Assessment of speech communication", ISO 9921:2003.
- [9] Hearing in the Communication Society. 2005. Speech recognition tests for different languages, FP6-004171 HEARCOM.
- [10] Fuchs, G. L. y J. Osuna . 1965. "Medición de inteligibilidad". *Memoria de las Primeras Jornadas Latinoamericanas de Acústica*. Universidad Nacional de Córdoba
- [11] Miñana, P. 1969. *Compendio práctico de acústica*. Barcelona: Labor.
- [12] Sommerhoff, J. y C. Rosas. 2007. "Evaluación de la inteligibilidad del habla en español". *Estudios Filológicos*
- [13] Aubanel, V., Lecumberri, M.L.G., Cooke, M. "The sharvard corpus: A phonemically-balanced Spanish sentence resource for audiology" *International Journal of Audiology* Volume 53, Issue 9, September 2014, 633-638
- [14] dr.ir. Emanu"el A.P. Habets. "Room Impulse Response Generator".September 20, 2010
- [15] The Castilian Spanish Hearing in Noise Test (*Huarte, 2008*)
- [16] Bamford-Kowal-Bank (BKB) (Bench y Bamford, 1979)

[17] Lena L.N. Wong, Elaine H.N. Ng “Characterization of speech understanding in various types of noise” (J. Acoustical Society of America 2012)

A. ANEXOS

a) Anexo A

Plan de actuación

Para realizar los test de forma eficiente, se ha hecho este plan de actuación. Con el fin de obtener un mínimo de 10 Test por ruido. Teniendo en cuenta que a cada encuestado se le va a realizar 3 test distintos se han creado 6 combinaciones distintas, para que a la misma persona no se le repita el mismo ruido.

Cada combinación será repetida 5 veces, obteniendo así la siguiente tabla. Que nos indica que necesitaremos de 30 encuestados.

SSN=ruido aleatorio.

Combinación A	Cafetería S	Tren	Carretera	5
Combinación B	Cafetería M	Carretera	SSN	5
Combinación C	Cafetería A	Tren	SSN	5
Combinación D	Cafetería S	Tren	SSN	5
Combinación E	Cafetería M	Tren	Carretera	5
Combinación F	Cafetería A	Carretera	SSN	5

Una vez realizado el siguiente plan, se observa que tenemos un total de 90 test.

Cafetería S	Cafetería M	Cafetería A	Tren	Carretera	SSN	
10	10	10	20	20	20	90

Además para la distribución de listas de frases se ha realizado la siguiente tabla. Donde por ejemplo, para el primer encuestado de cada combinación se les pasara las listas 1,2 y 3.

Encuestado/Combinación	1	2	3	4	5
Combinación A	123	451	234	512	345
Combinación B	123	451	234	512	345
Combinación C	123	451	234	512	345
Combinación D	123	451	234	512	345
Combinación E	123	451	234	512	345
Combinación F	123	451	234	512	345

Por ultimo para decidir el género de la voz que reproduce las frases se ha hecho la siguiente tabla donde para los primeros encuestados de cada combinación, la primera lista es en Femenino, la segunda en Masculino y la tercera en femenino.

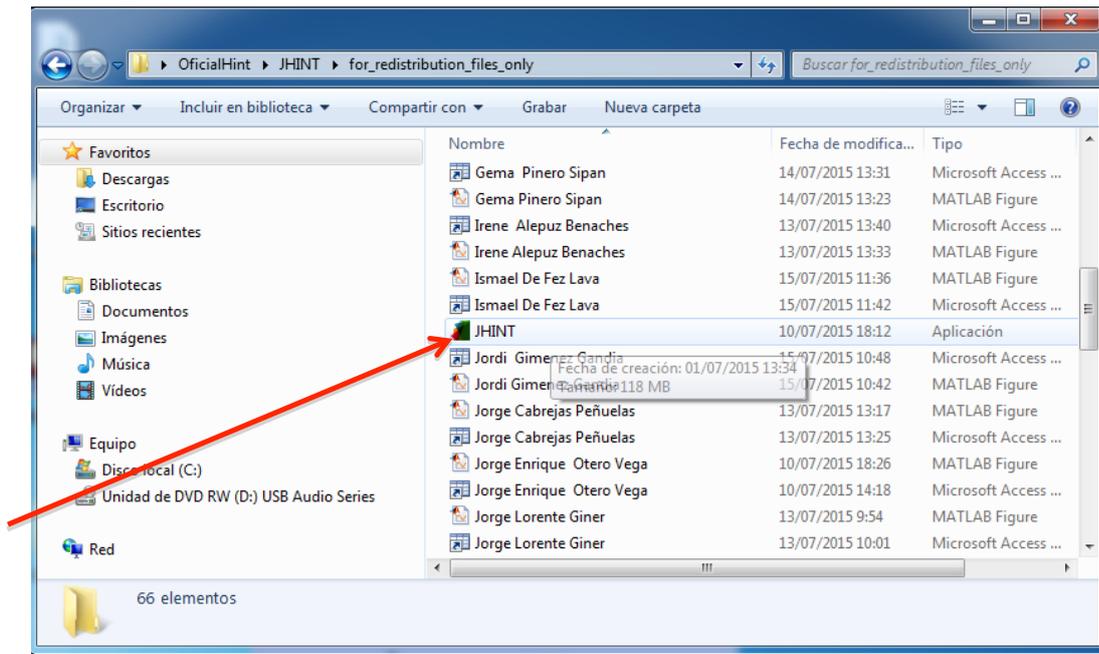
1	2	3	4	5
fmf	mfm	ffm	mmf	fmf

Una vez realizado el plan de actuación, ya se puede empezar a realizar los test.

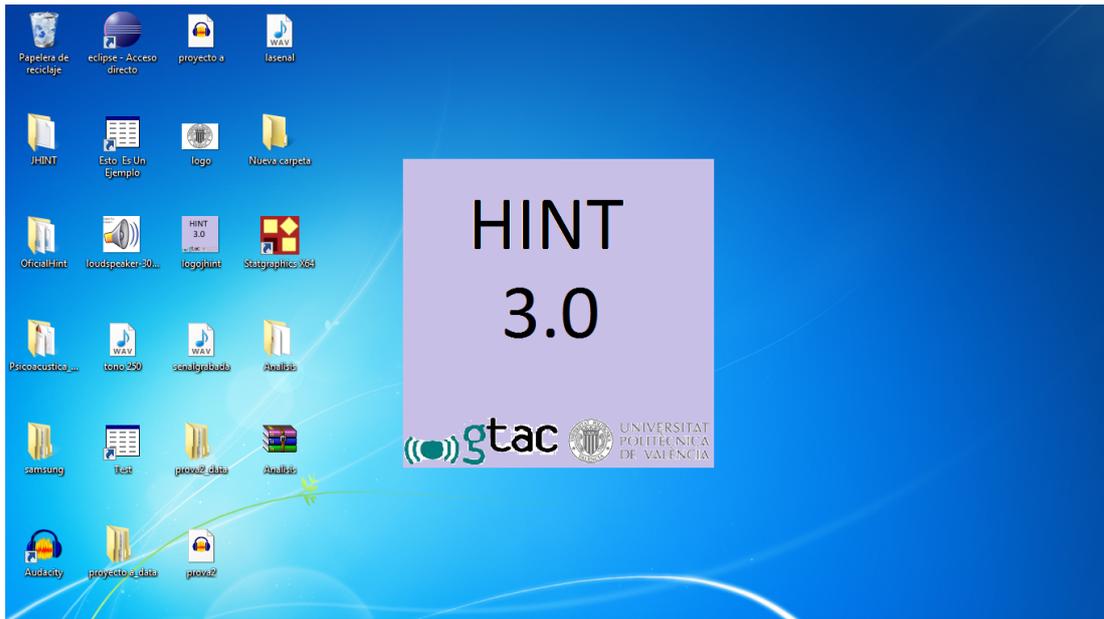
b) Anexo B

Este es un tutorial de la aplicación JHINT esta es una aplicación que realiza el test HINT (Hearing In Noise Test).

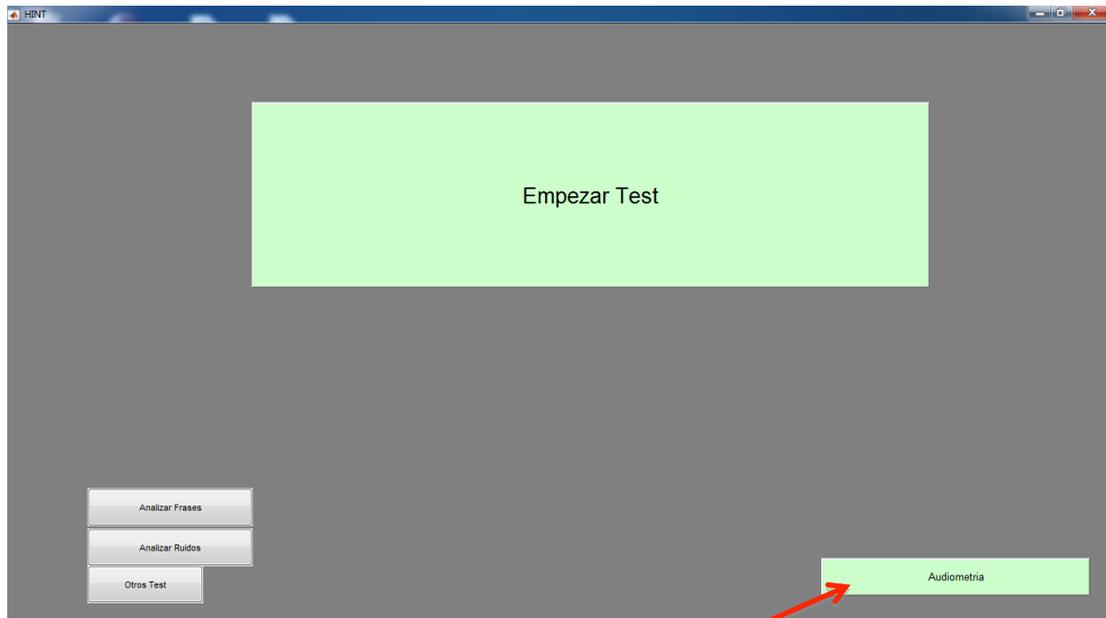
El Primer paso es acceder al ejecutable, con nombre HINT, y darle al doble click.



Entonces se ejecutara la aplicación JHINT en la versión 3.0 este ofrece la siguiente imagen mientras carga el programa.



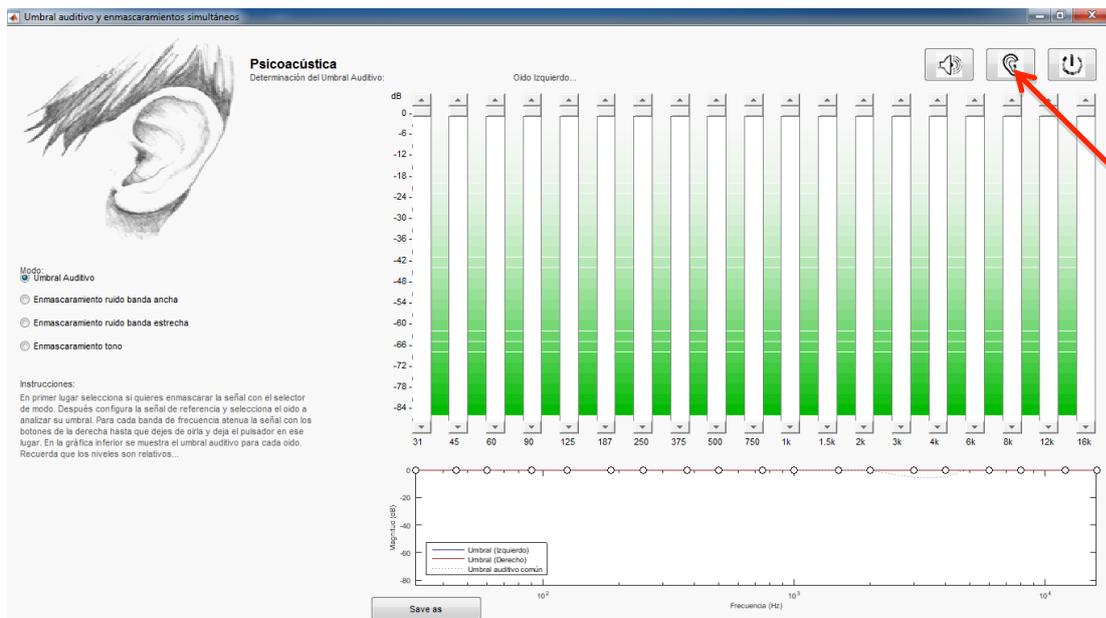
Una vez cargado el programa, aparece la página principal de la aplicación. Y el primer paso es realizar una audiometría, para ello solo se necesita pulsar el botón de AUDIOMETRIA, con este accederemos a un programa que calcula el umbral de audición del encuestado.



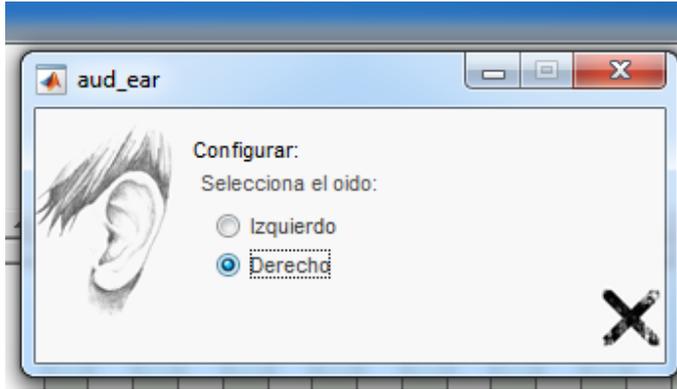
Una vez abierto el programa Psicoacústica que realiza la audiometría hay que seleccionar el punto 1, Umbral auditivo y enmascaramiento, y seleccionar el botón que indica la flecha en la imagen, para pasar al siguiente punto.



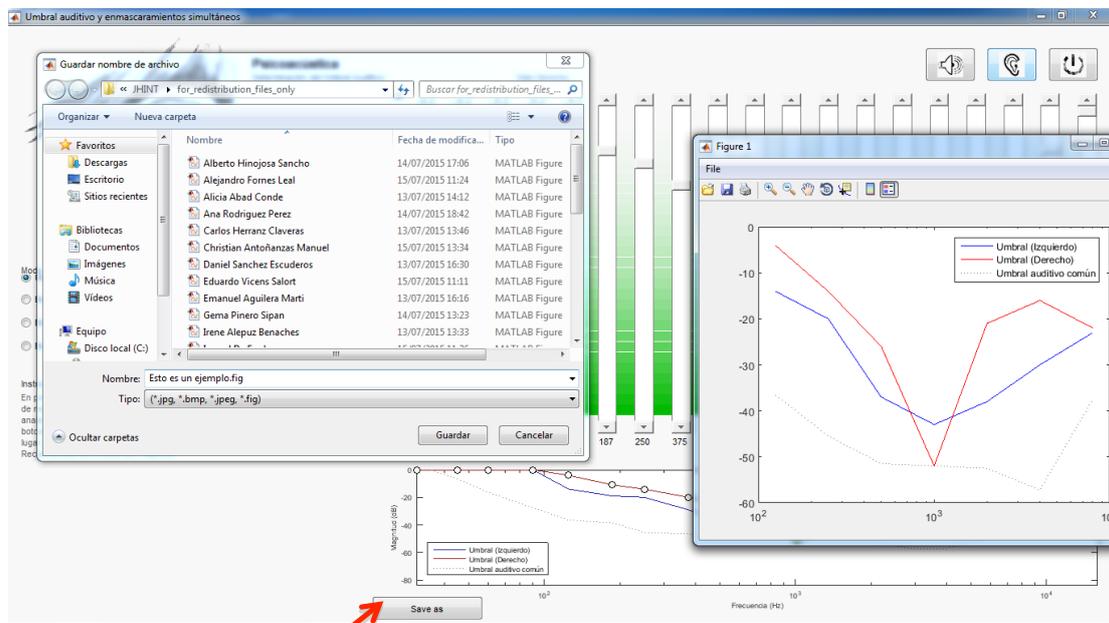
Este abrirá la siguiente ventana, en ella se calculara el umbral de audición para las siguientes frecuencias, 125,250,500,1000,2000,4000 y 8000 Hz. Solo con realizar estas, al guardar el programa interpolara y creara el grafico entero. Pulsando las flechas de arriba de la frecuencia, se irá disminuyendo el nivel de presión en el que sonara el tono por los auriculares, en pasos de 3 dB hasta que el usuario, no oiga el tono. El procedimiento es empezar siempre por el de 1000 Hz y luego ir alternando las frecuencias, además la barra puede deslizarse así que en algunas se empezara desde la no audición de la frecuencia, y se irá subiendo. Una vez calculado el umbral de audición del oído izquierdo, se va a pulsar lo que indica la flecha roja, para cambiar al otro oído.



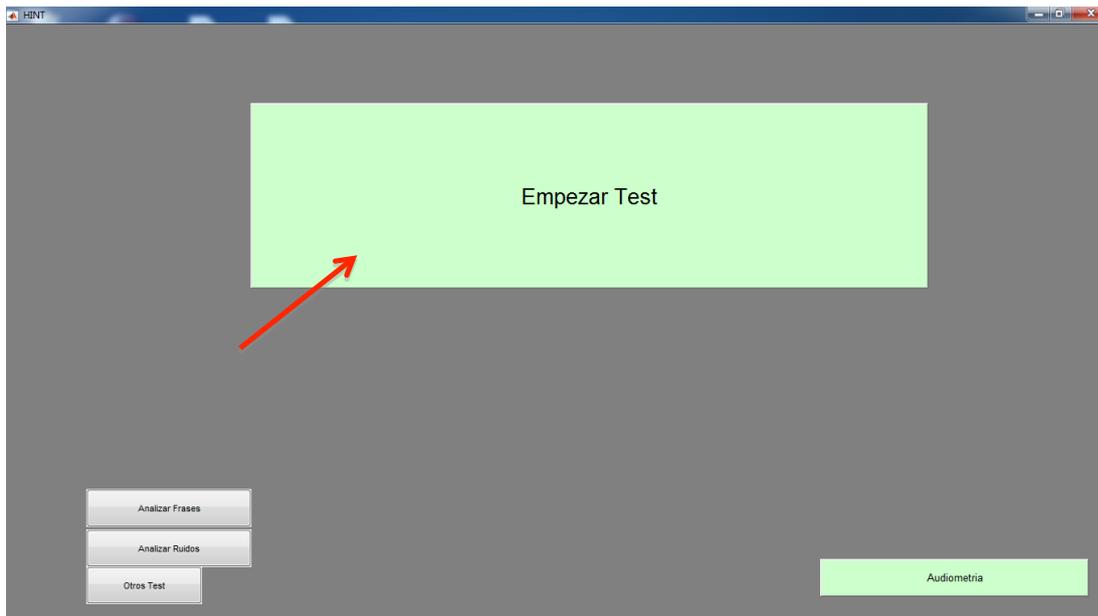
Entonces aparece otra pantalla que nos deja cambiar al oído derecho y se procederá a realizar el mismo procedimiento con el oído derecho.



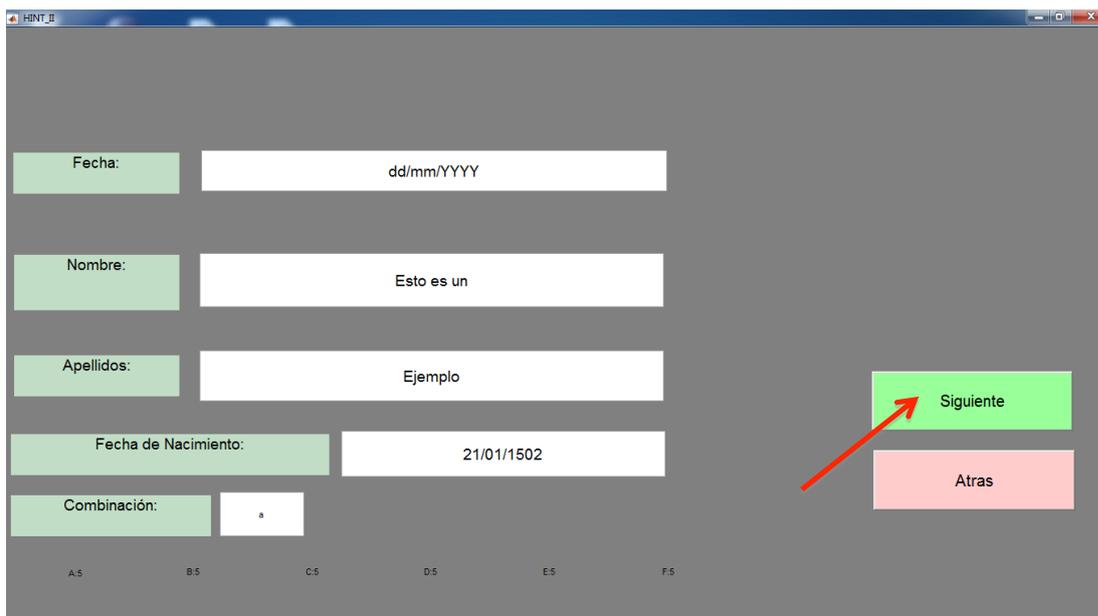
Una vez realizada la audiometría para el oído derecho se clicara el botón SAVE AS, en este momento hay que guardar la audiometría, con el nombre del encuestado y añadiendo “.fig”



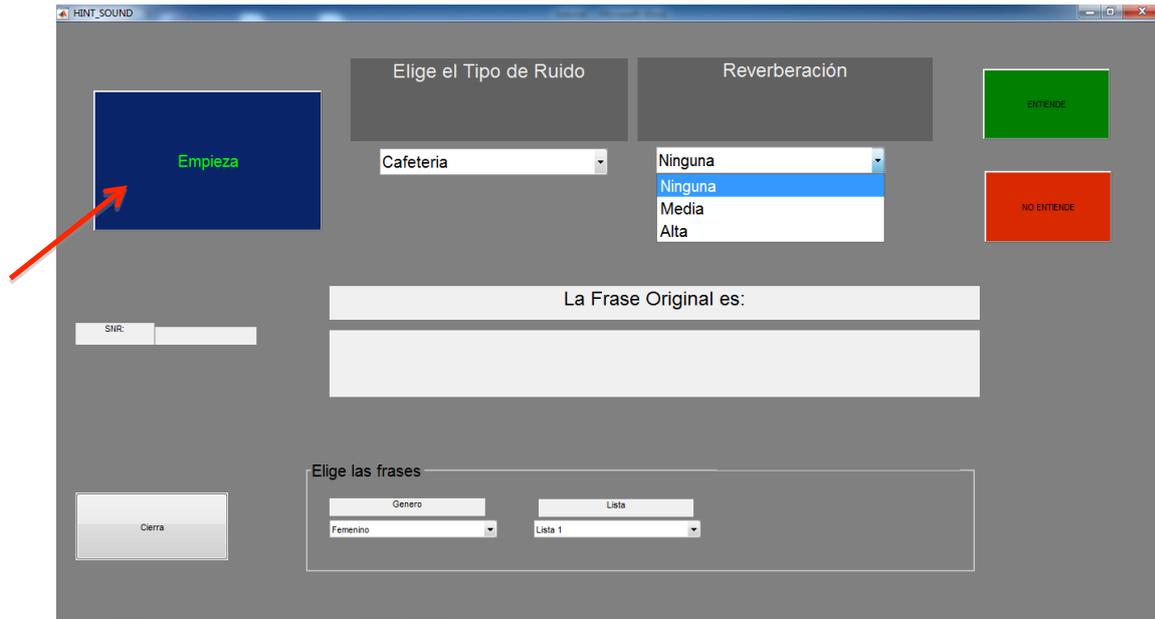
Una vez realizada la audiometría volvemos a la página principal, y procederemos a realizar el test clicando en el botón “Empezar Test”.



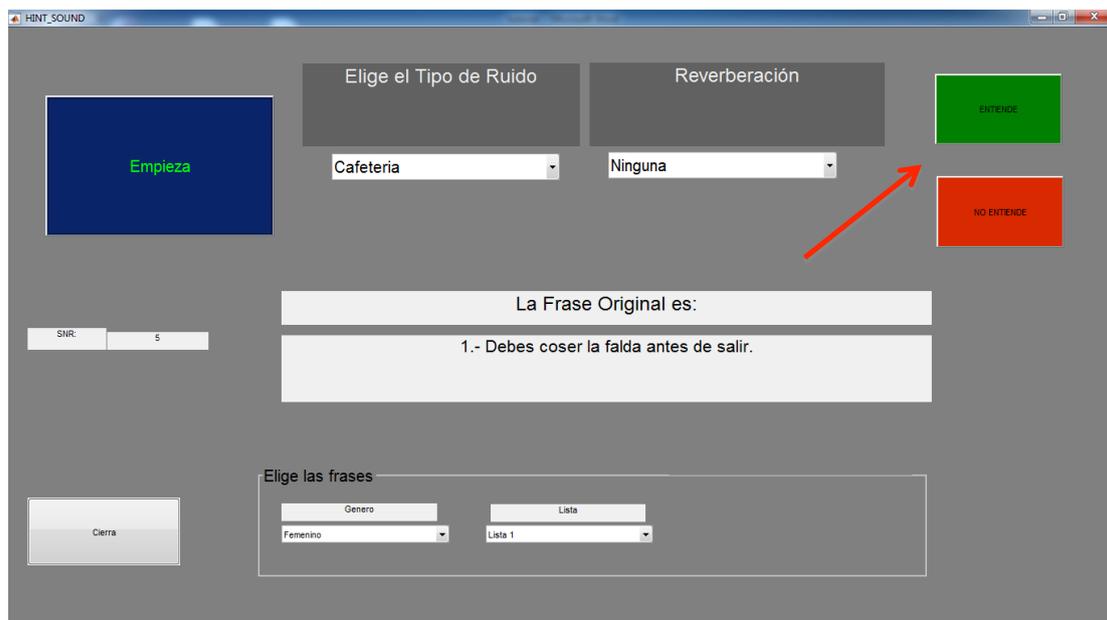
Esto nos llevara a la página para recoger datos del encuestado, en esta se pondrá la fecha de la encuesta en el siguiente formato, “dd/mm/YYYY”, el nombre del encuestado, los apellidos del mismo, y la fecha de nacimiento de este, no será necesario rellenar el campo de combinación, ya que este se utilizó para realizar este proyecto y solo sirve para llevar un recuento. Una vez rellenados los campos se deberá pulsar el botón siguiente.



Esto nos llevara a la siguiente página de la aplicación en este punto se debe seleccionar el ruido, la reverberación *solo en el caso de que se escoja el ruido de cafetería, el género de la voz que pronunciara las frases y por último la lista a reproducir. El siguiente paso es pulsar el botón de empieza. Para cada encuestado se deberán realizar 3 test con sus respectivos ruidos y sus respectivas listas de frases a reproducir.



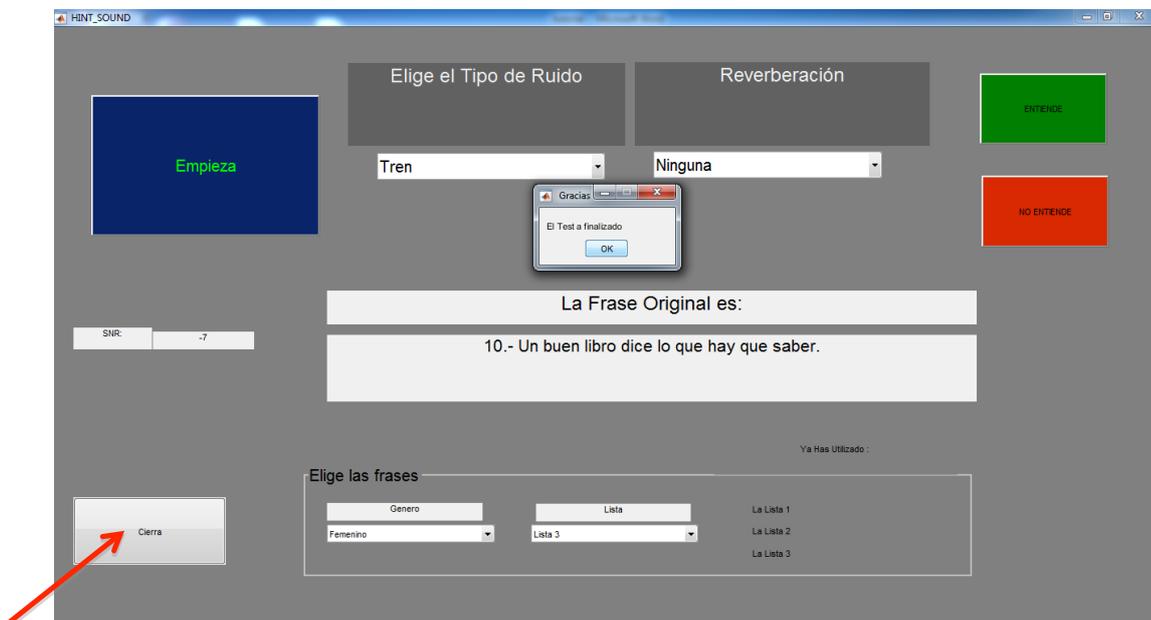
Al poner empieza, se reproduce la primera frase del test, el usuario encuestado deberá repetir la frase que escucha, y repetirla, en el caso de que falle una palabra se deberá pulsar el botón, no entiende, y en el caso de que acierte todas las palabras se deberá pulsar el botón de entiende, esto durante las siguientes 10 frases. Los artículos y tiempos verbales no penalizarán.



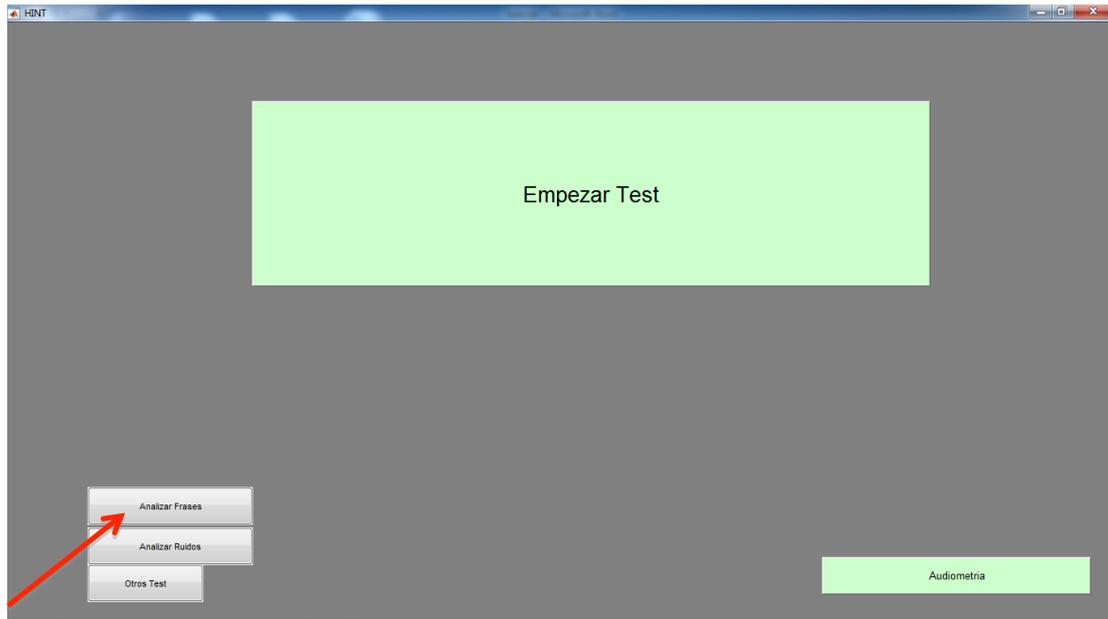
Una vez se llega a la última frase, el test nos avisa de que ha finalizado el test. Además nos la aplicación nos indica que frases han sido utilizadas. Por lo que el siguiente paso es utilizar otra lista y otro ruido y realizar otro test.



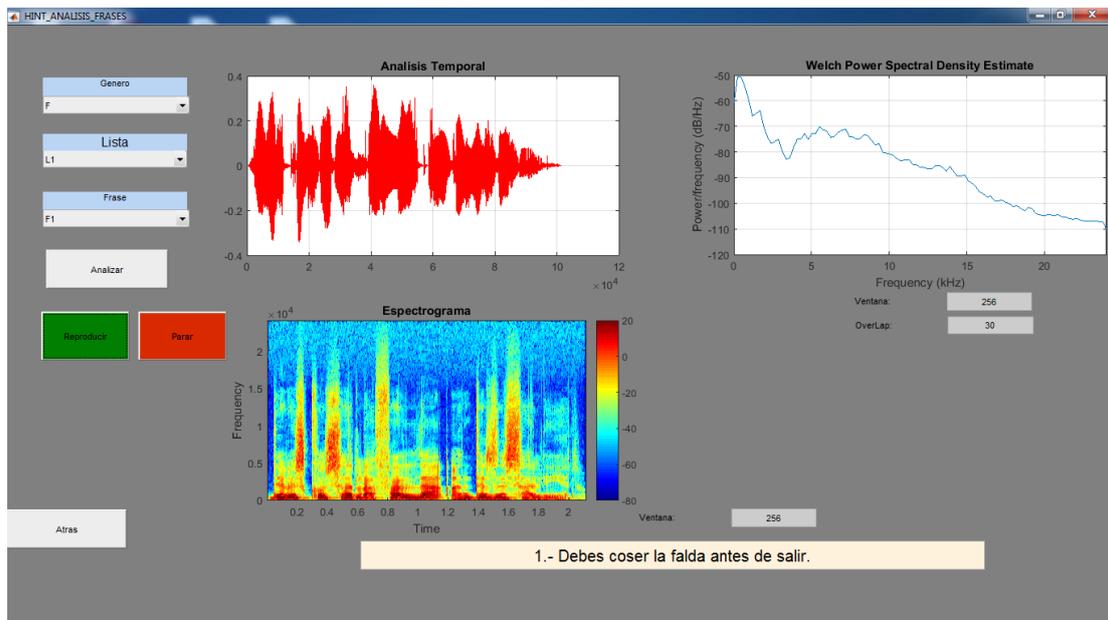
Cuando tengamos los 3 test realizados, pulsaremos el botón cierra, en este punto el test estará finalizado y pasaremos a la página principal.



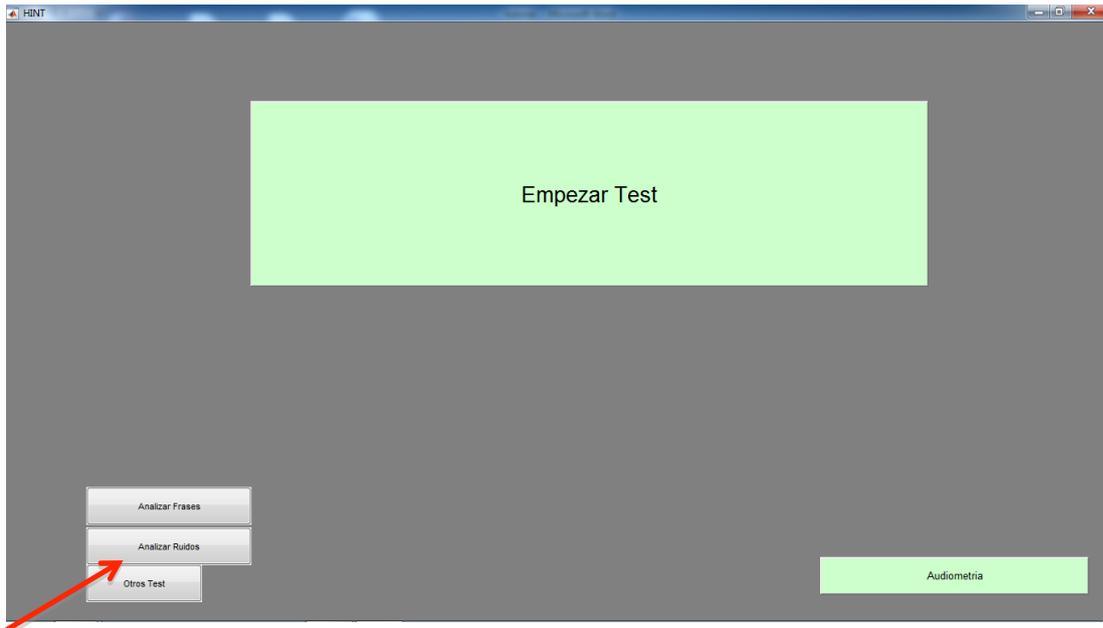
Ahora veremos las prestaciones que nos da la aplicación. En la parte inferior izquierda podemos analizar tanto las frases, los ruidos como los test. Primero vamos a ver que nos permite analizar de las frases. Pulsando el botón analiza frases.



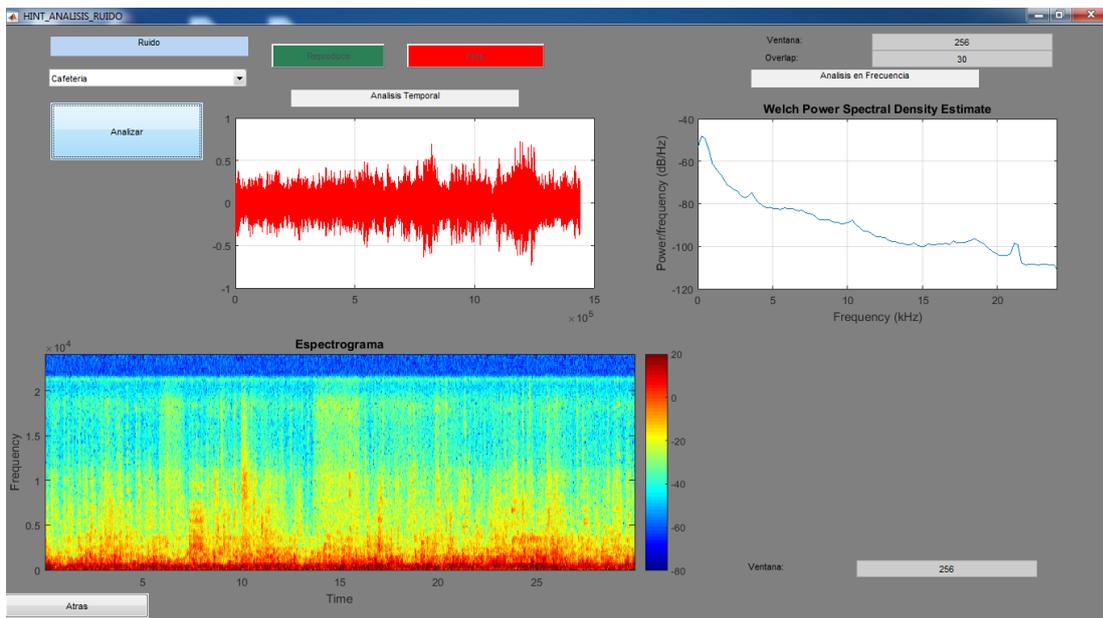
Este botón nos llevara a la siguiente página donde podemos elegir el género, la lista y la frase, una vez elegido se deberá pulsar analizar. Y aparece un gráfico de la potencia versus tiempo, otro con el análisis espectral y otro con un espectrograma, además se puede modificar la ventana y el solapamiento, para ver con más precisión los gráficos. Además está la opción de reproducir la frase.



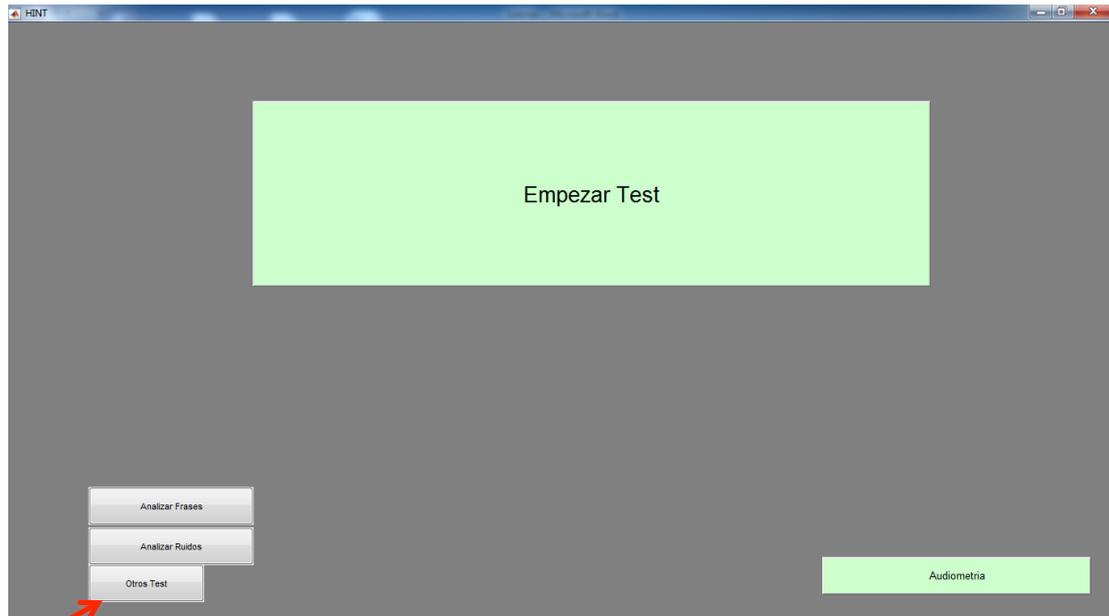
Si decidimos analizar la opción de ruido, pulsando analizar ruido en la pagina principal.



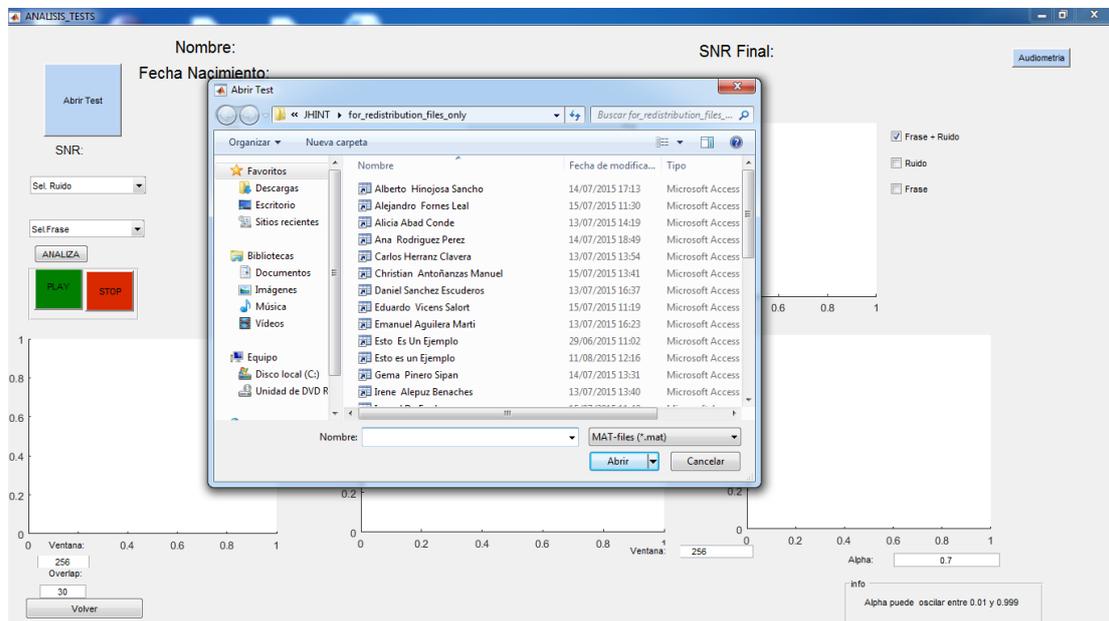
Esto nos da la opción de analizar todos los ruidos, primero se selecciona el ruido, después se pulsa el botón analizar, aquí tenemos las mismas prestaciones que en las frases, análisis temporal, análisis en frecuencia y espectrograma. Aquí podemos modificar la ventana y el solapamiento para tener más precisión. Además de poder reproducir los ruidos.



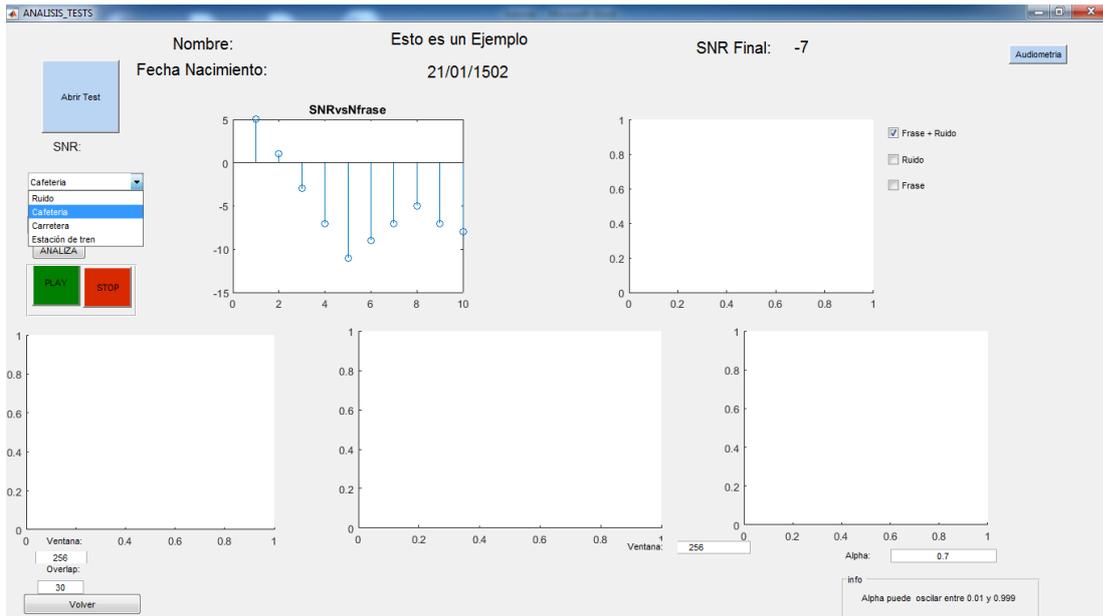
si volvemos a la pagina principal, tendremos la posibilidad de analizar los test realizados.



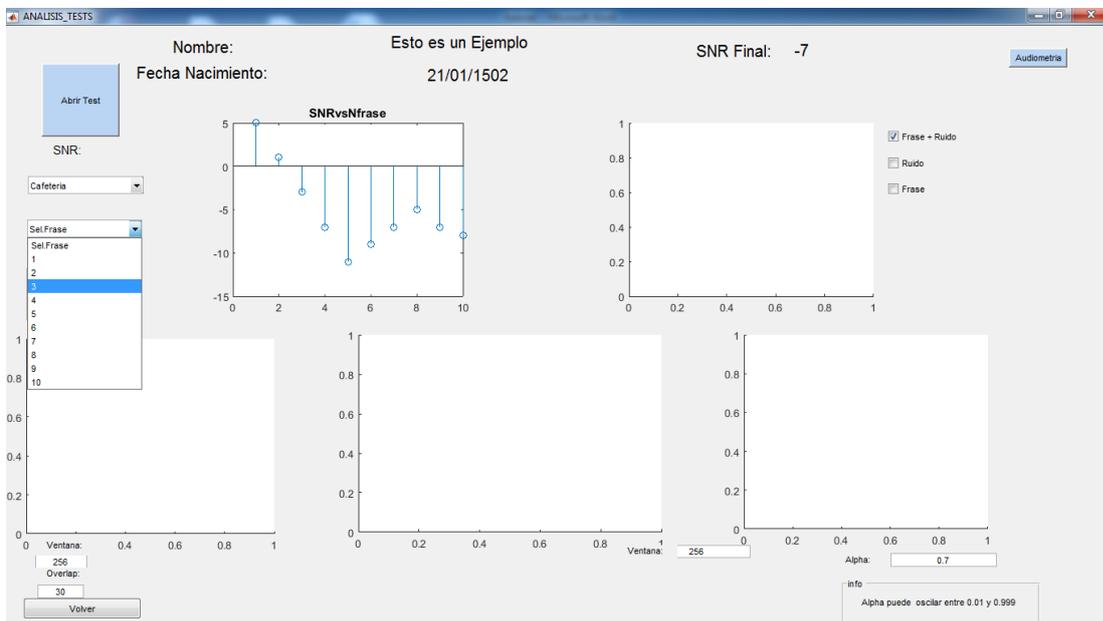
Si pulsamos en abrir test nos ofrece la opción de seleccionar el test que deseamos analizar.



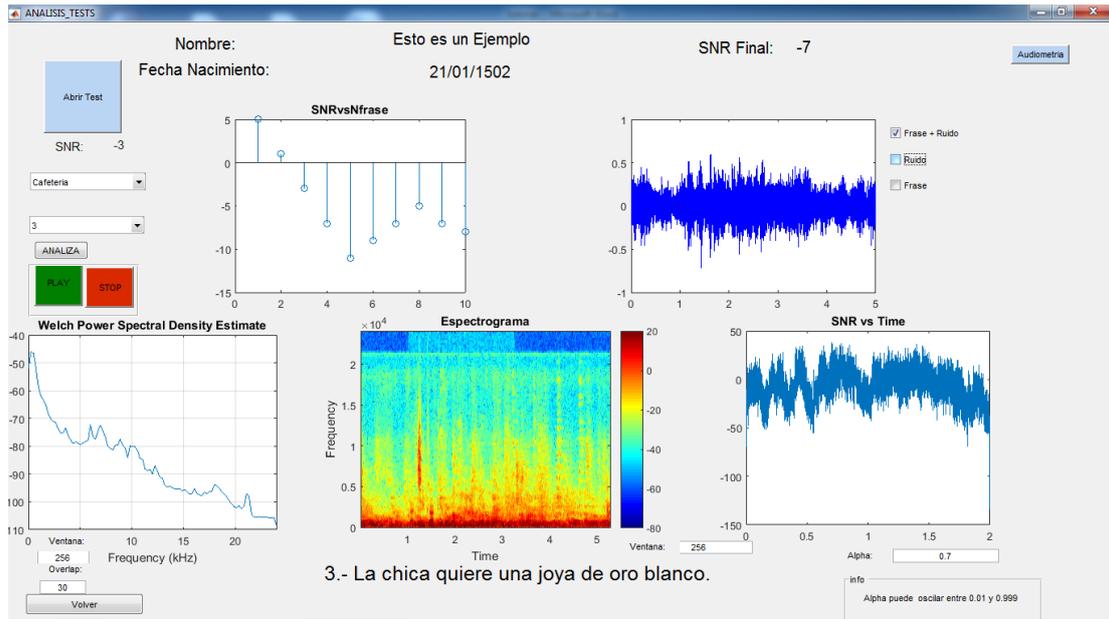
Una vez seleccionado aparece el nombre del encuestado su fecha de nacimiento y la relación señal a ruido final. Aquí tienes la opción de elegir los ruidos y una vez seleccionado el ruido, puedes seleccionar una de las 10 frases. En el primer gráfico aparece la SNR versus frase, donde vemos cada frase con cada relación señal a ruido.



En este punto se elige la frase a analizar.

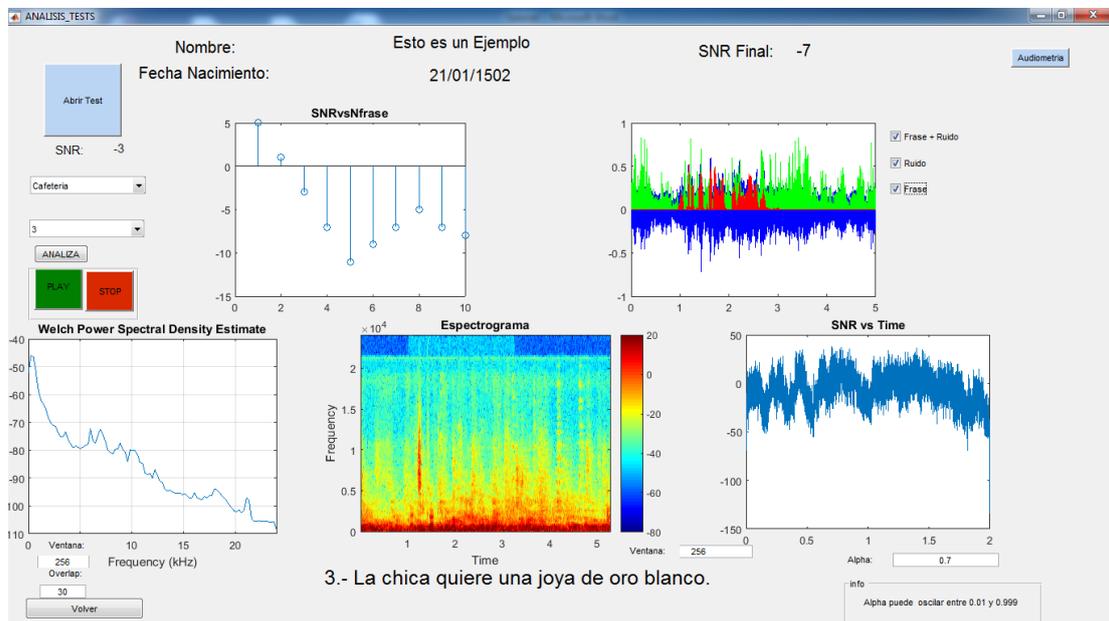


Una vez elegida la frase tenemos en el segundo grafico un análisis temporal, en el tercero tenemos un análisis espectral, luego el espectrograma y por último la relación señal a ruido versus tiempo.



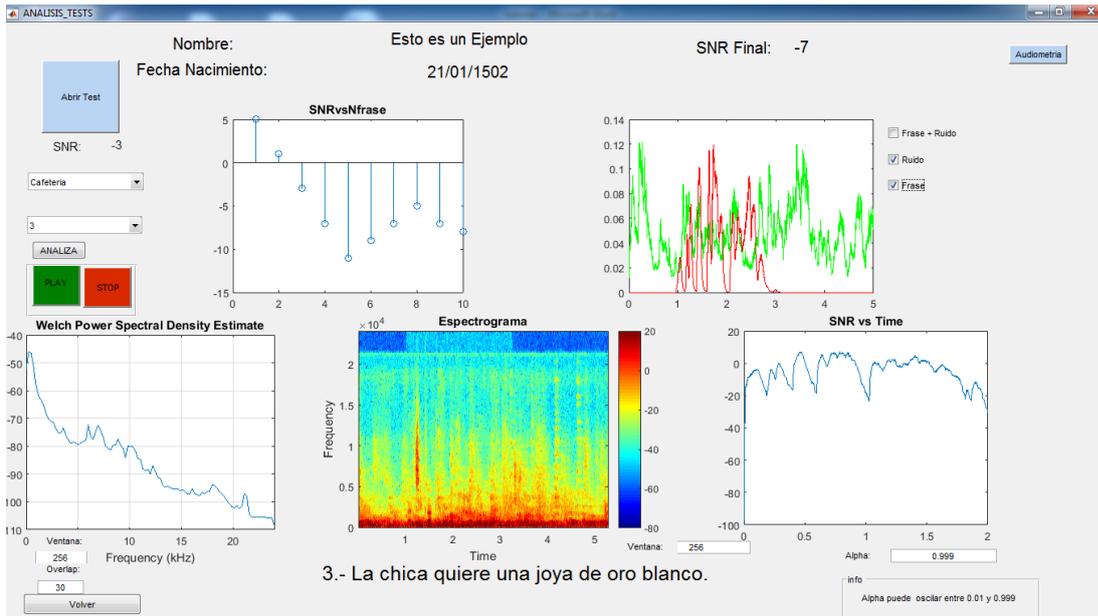
3.- La chica quiere una joya de oro blanco.

Si pulsamos las pestañas ruido y frase podremos ver la potencia de estas.



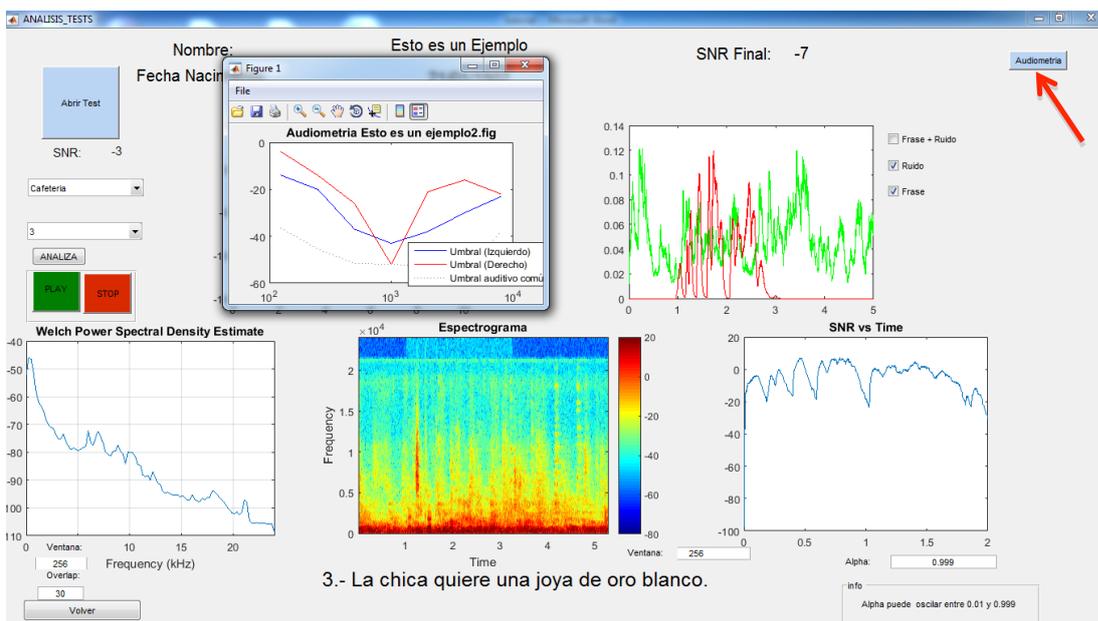
3.- La chica quiere una joya de oro blanco.

Si quitamos la pestaña de frase+ruido, solo observaremos las potencias del ruido y la frase. Además modificando el alfa podemos ver con más precisión la SNR vs Tiempo.



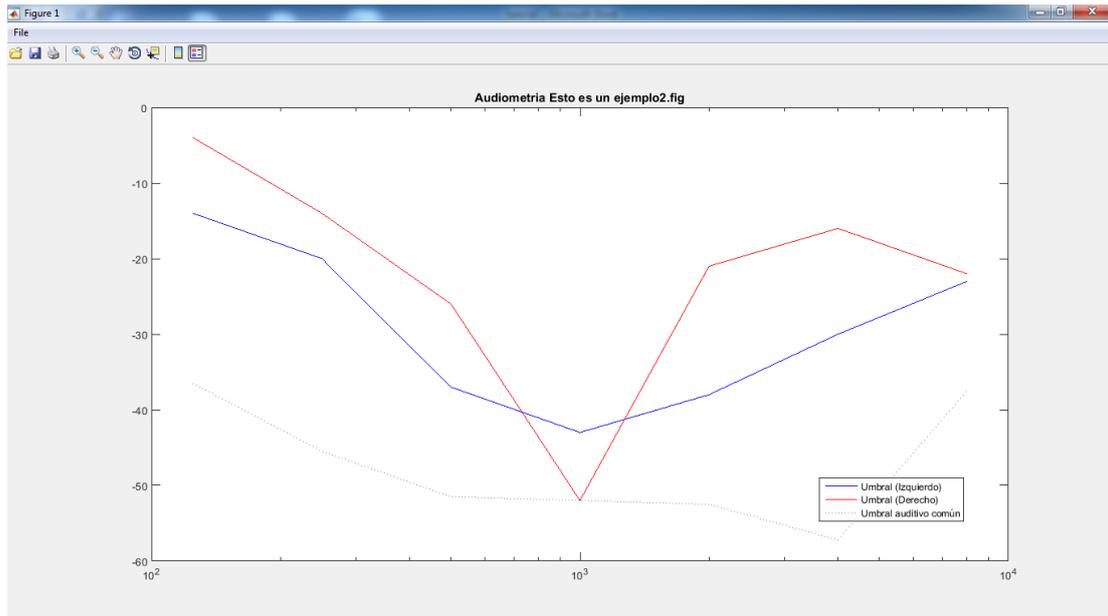
3.- La chica quiere una joya de oro blanco.

Por ultimo pulsando en audiometría, accedemos a la audiometría del encuestado, solo hay que seleccionar el nombre del encuestado y observaremos la gráfica del umbral de audición.



3.- La chica quiere una joya de oro blanco.

En el siguiente grafico vemos la audiometría.



Con estos sencillos pasos tenemos un tutorial para poder controlar la aplicación JHINT.

c) Anexo C

El código Matlab lo podemos encontrar en la carpeta de ANEXOS con el nombre de Código Matlab

d) Anexo D

LISTA DE FRASES BALANCEADAS FONETICAMENTE

Lista 1 (Lista 6)

- 1.- Debes coser la falda antes de salir.
- 2.- Se usa una sierra radial para cortar tablas.
- 3.- La chica quiere una joya de oro blanco
- 4.- En esta zona, el aire es muy puro
- 5.- Metió un clavo en la tabla para tapan el hueco
- 6.- Ayer vino un nuevo huésped al hotel
- 7.- La huella de sangre mostró el rastro al hombre.
- 8.- La mujer se pudo poner de pie otra vez
- 9.- Tomó la ruta más larga a su casa
- 10.- La chica del cine vendió cinco bonos.

Lista 2 (Lista 17)

- 1.- La risa del niño nos cambió el mal humor.
- 2.- Hay poco papel, por lo que no lo gasto
- 3.- Una chispa quemó el mantel justo en el medio
- 4.- La falda roja se secó en una percha fina.
- 5.- Todos los niños se ríen cuando van al circo.
- 6.- Es mejor una silla alta para ver el tenis.
- 7.- Las peras verdes están aún duras.
- 8.- El nuevo doctor tuvo un grave error.
- 9.- Lanzó la paja a través de la puerta de la cuadra.
- 10.- El lago se hiela siempre que hace frío.

Lista 3 (Lista 40)

- 1.- Deje su vaso dentro de la bolsa negra
- 2.- Hay una concha de color rosa en la playa.
- 3.- Pinta las puertas de rojo y la pared gris.
- 4.- Un rey mandó en la nación veinte años.
- 5.- La mujer tiene mejor vista que tacto.
- 6.-Una forma de ser rico es no gastar mucho.
- 7.- Sumar de prisa causó un error en la cuenta.
- 8.- Oyó un coche llegar a la puerta de casa
- 9.- Los tacos se hunden fácil en el césped blando.
- 10.- Un buen libro dice lo que hay que saber.

Lista 4 (Lista 43)

- 1.- Es mejor pensar un poco antes de hablar.
- 2.- Siente el calor de esa débil llama.
- 3.- El pastor daba pienso a sus cerdos para la cría.
- 4.- Coge la sierra por el mango para hacer fuerza.
- 5.- El chico le dio cinco euros a un pobre.
- 6.- Las moscas zumban con el sol y el calor de junio.
- 7.- Dale la mano a este niño rubio.
- 8.- El grupo de gente se sumó a la fuerte lucha.
- 9.- Vierta el guiso de la olla al plato grande.
- 10.- Salta la valla y vete hacia el banco.

Lista 5 (Lista 59)

- 1.- Marca el solar con un cartel con letras en rojo.
- 2.- El potro saltó muy bien la valla.
- 3.- El ciclo de la vida tiene varias fases.
- 4.- El humo sale de todas las grietas de la casa.
- 5.- La planta creció mucho detrás del jardín.
- 6.- Quemó la ropa de lana con una chispa de fuego.
- 7.- El hombre cojo anda por medio de la vía.
- 8.- Siempre nos gusta tener buen tiempo.
- 9.- El gran foco de neón rojo se cayó.
- 10.- El gran sable se hizo con bronce puro.

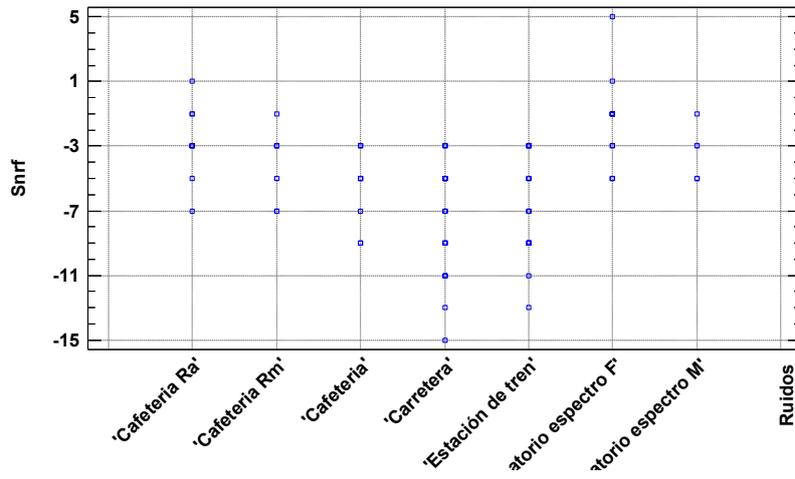
e) Anexo E**Resultados**

Aquí podemos ver un análisis exhaustivo de los resultados, primero de Snr final Frente a los distintos ruidos.

Resumen Estadístico para Snrf

Ruidos	Recuento	Promedio	Mediana	Moda	Varianza	Desviación Estándar
'Cafetería Ra'	10	-2,8	-3,0	-3,0	4,8444	2,201
'Cafetería Rm'	10	-4,4	-4,0	-3,0	4,4889	2,1187
'Cafetería'	10	-5,4	-5,0	-5,0	5,1556	2,2706
'Carretera'	20	-7,4	-7,0	-5,0	12,042	3,4702
'Estación de tren'	20	-6,9	-7,0	-9,0	9,0421	3,007
'Ruido aleatorio espectro F'	13	-1,6154	-1,0	-1,0	7,5897	2,7549
'Ruido aleatorio espectro M'	7	-3,8571	-5,0	-5,0	2,4762	1,5736
Total	90	-5,1111	-5,0	-3,0	11,628	3,41

Dispersión por Código de Nivel



Medias y 90,0% de Fisher LSD

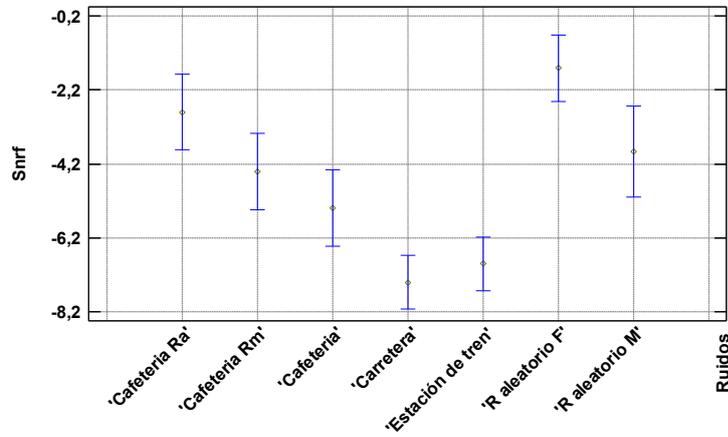


Gráfico Caja y Bigotes

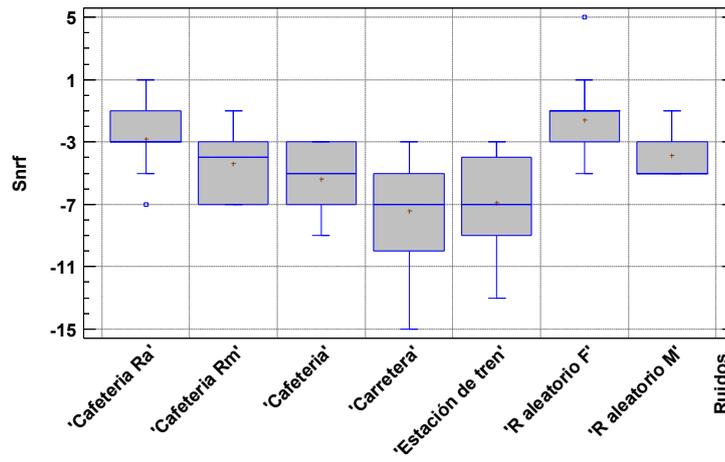


Gráfico de Medianas con Intervalos del 95,0% de Confianza

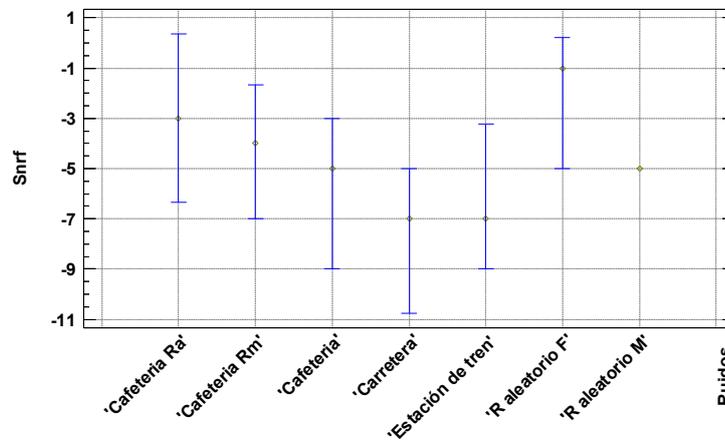
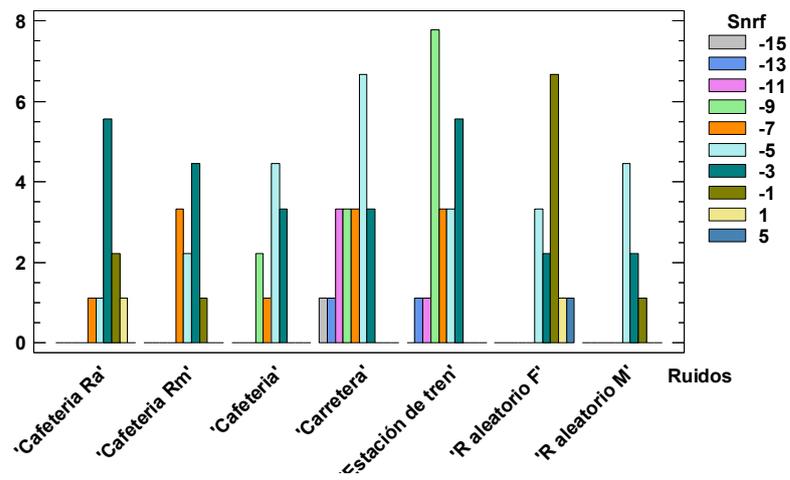


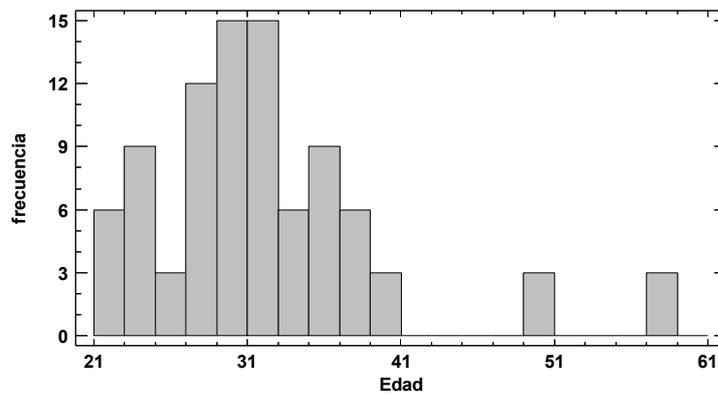
Diagrama de Barras para Ruidos según Snrf



Edad

Aquí podemos ver un análisis estadístico de las edades.

Histograma



Resumen Estadístico para Edad

Recuento	90
Promedio	32,7
Mediana	31,5
Moda	
Varianza	55,628
Desviación Estándar	7,4584
Coefficiente de Variación	22,809%
Error Estándar	0,78619
Mínimo	23,0
Máximo	58,0

Rango	35,0
Cuartil Inferior	29,0
Cuartil Superior	36,0
Sesgo Estandarizado	5,9396
Curtosis Estandarizada	6,2232
Suma de Cuadrados	1,0119E5

Por listas

Aquí tenemos un análisis estadístico de Snr final Frente a las distintas listas.

Resumen Estadístico para Snrf

Listas	Recuento	Promedio	Mediana	Moda	Varianza	Desviación Estándar	Coficiente de Variación
'FLista1'	6	-3,6667	-3,0	-3,0	2,6667	1,633	-44,536%
'FLista2'	11	-4,6364	-5,0	-5,0	18,255	4,2725	-92,153%
'FLista3'	18	-5,7778	-5,0	-5,0	10,418	3,2277	-55,865%
'FLista4'	1	5,0	5,0	5,0			%
'FLista5'	12	-4,1667	-3,0	-3,0	9,7879	3,1286	-75,085%
'MLista1'	11	-6,8182	-5,0	-5,0	4,3636	2,0889	-30,638%
'MLista2'	7	-7,5714	-9,0	-9,0	18,286	4,2762	-56,478%
'MLista3'	1	-13,0	-13,0	-13,0			%
'MLista4'	17	-3,9412	-3,0	-3,0	5,0588	2,2492	-57,069%
'MLista5'	6	-5,0	-5,0	-5,0	4,8	2,1909	-43,818%
Total	90	-5,1111	-5,0	-3,0	11,628	3,41	-66,717%

Medias y 95,0% de Fisher LSD

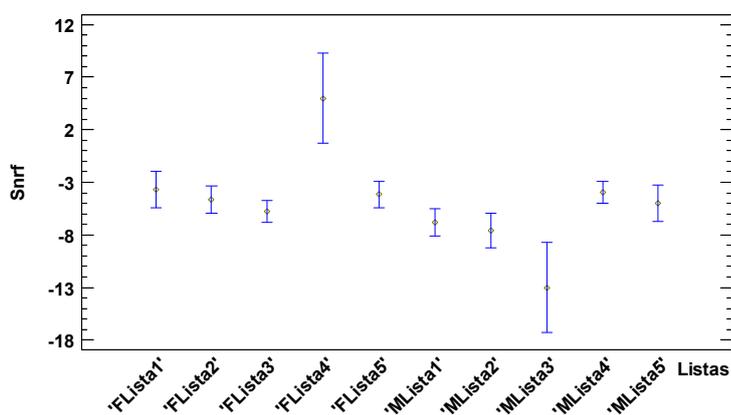
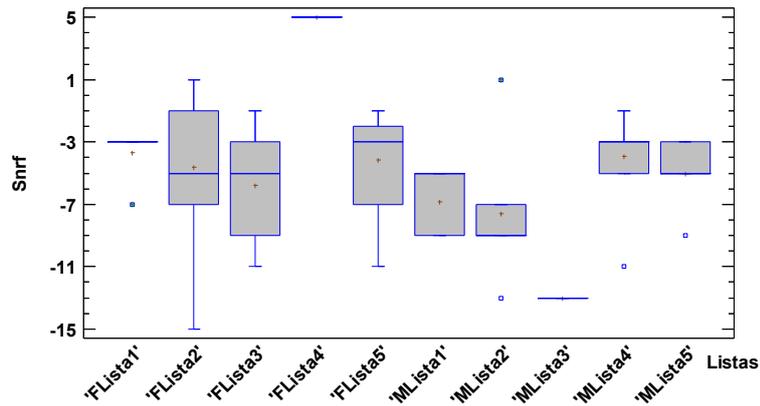


Gráfico Caja y Bigotes

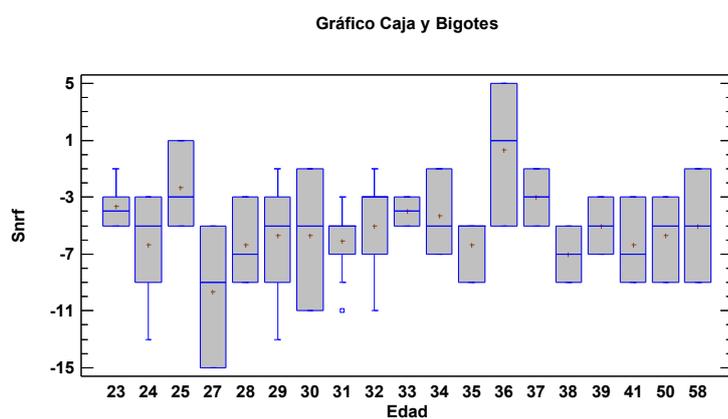
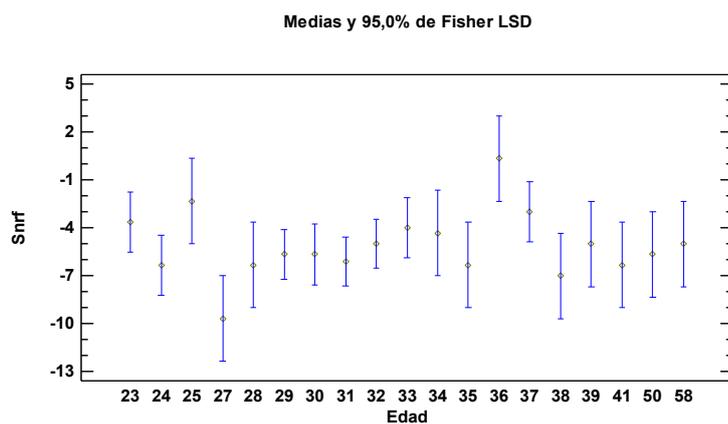


Por edad

Aquí tenemos un análisis de Snr final Frente a las edades.

Resumen Estadístico para Snrf

Edad	Recuento	Promedio	Mediana	Moda	Varianza	Desviación Estándar	Coefficiente de Variación
23	6	-3,6667	-4,0	-5,0	2,6667	1,633	-44,536%
24	6	-6,3333	-5,0	-3,0	17,067	4,1312	-65,229%
25	3	-2,3333	-3,0		9,3333	3,0551	-130,93%
27	3	-9,6667	-9,0		25,333	5,0332	-52,068%
28	3	-6,3333	-7,0		9,3333	3,0551	-48,238%
29	9	-5,6667	-5,0		17,0	4,1231	-72,761%
30	6	-5,6667	-5,0		21,867	4,6762	-82,521%
31	9	-6,1111	-5,0	-5,0	7,1111	2,6667	-43,636%
32	9	-5,0	-3,0	-3,0	11,0	3,3166	-66,332%
33	6	-4,0	-4,0		1,2	1,0954	-27,386%
34	3	-4,3333	-5,0		9,3333	3,0551	-70,501%
35	3	-6,3333	-5,0	-5,0	5,3333	2,3094	-36,464%
36	3	0,33333	1,0		25,333	5,0332	1510,0%
37	6	-3,0	-3,0		3,2	1,7889	-59,628%
38	3	-7,0	-7,0		4,0	2,0	-28,571%
39	3	-5,0	-5,0		4,0	2,0	-40,0%
41	3	-6,3333	-7,0		9,3333	3,0551	-48,238%
50	3	-5,6667	-5,0		9,3333	3,0551	-53,913%
58	3	-5,0	-5,0		16,0	4,0	-80,0%
Total	90	-5,1111	-5,0	-3,0	11,628	3,41	-66,717%



Por genero

Y por ultimo un análisis estadístico de el genero de las frases reproducidas frente la Snr final.

Resumen Estadístico para Snrf

Genero	Recuento	Promedio	Mediana	Moda	Varianza	Desviación Estándar	Coficiente de Variación
'F'	48	-4,625	-4,0	-3,0	12,878	3,5885	-77,59%
'M'	42	-5,6667	-5,0	-5,0	9,8862	3,1442	-55,486%
Total	90	-5,1111	-5,0	-3,0	11,628	3,41	-66,717%

Medias y 95,0% de Fisher LSD

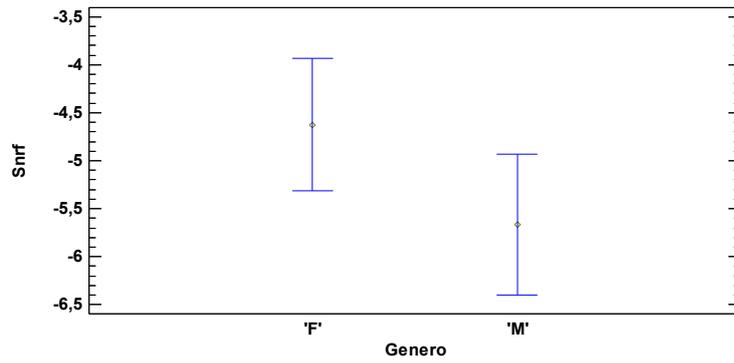


Gráfico Caja y Bigotes

