

## DESARROLLO DE HERRAMIENTA SOFTWARE PARA LA REALIZACIÓN DE TEST DE AUDICIÓN CON RUIDO “JHINT”

PACS:43.60.Qv

Rodríguez, Josep; Ferrer, Miguel; Piñero, Gema; de Diego, María  
Inst. de Telecom. Y Aplicaciones Multimedia (iTEAM)  
Univ. Politécnica de Valencia (UPV)  
Camino de Vera, s/n.  
46022 Valencia  
Tel: +34963877007, ext. 73008,  
Fax: +34963877309,  
E-mail: jorodsen@epsg.upv.es

### 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. The HINT test has emulated with different noises (road, Café, Random noise ...) to calculate the SNR with 50% intelligibility. 50 sentences were used and evaluated 30 individuals without hearing impairments with ages between 23 and 58, in order to obtain an average result of the SNR ("Signal to Noise Ratio") for different noises.

### RESUMEN

Hearing In Noise Test (HINT o 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 con distintos ruidos (carretera, cafetería, ruido aleatorio...) para calcular el 50% de inteligibilidad. Se han utilizando 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") para los distintos ruidos.

### 1. INTRODUCCIÓN

Las deficiencias auditivas tratan de subsanarse mediante 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.

En este trabajo se presenta una herramienta que sirve para calcular la SNR (Signal to Noise Ratio) que corresponde al 50% de inteligibilidad, extrapolando la metodología descrita por Nilsson (1994) [1].

Dicha herramienta se ha llamado “JHINT” y mejora el test HINT original añadiendo la posibilidad de realizar el análisis para distintos ruidos procedentes de la realidad cotidiana, ya

que, en el Hint original se utiliza sólo ruido blanco o balbuceos. Además, incorpora un Corpus español extraído del Sharvard Corpus [2].

Existen varias formas de calcular la inteligibilidad, como por ejemplo el SRT (Speech reception threshold) [3], que es un procedimiento audiométrico de discurso no invasivo que mide el grado de pérdida de la audición para el habla. El umbral de reconocimiento de voz es el nivel de audiencia 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. El SRT se define como el nivel de audición más bajo a la que el encuestado 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 el SRT no es el mejor método, Por este motivo la aplicación HINT que utiliza frases cotidianas es un método mas efectivo, además, en la herramienta realizada JHINT se obtiene mayor efectividad al utilizar distintos ruidos como: ruido de carretera, ruido de cafetería, ruido de estación de tren y ruido aleatorio filtrado con el espectro de la voz que reproduce las frases.

La herramienta ha sido desarrollada en una aplicación mediante *GUI* (Graphical User Interface) de MATLAB, implementando el procedimiento descrito en el test Hint y que se detallará a continuación.

## 2. DESCRIPCIÓN DEL JHINT

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 [4].

El HINT es un mecanismo para el reconocimiento de voz adaptativo. 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 de la lista. Por el contrario, el nivel de presentación del ruido 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).

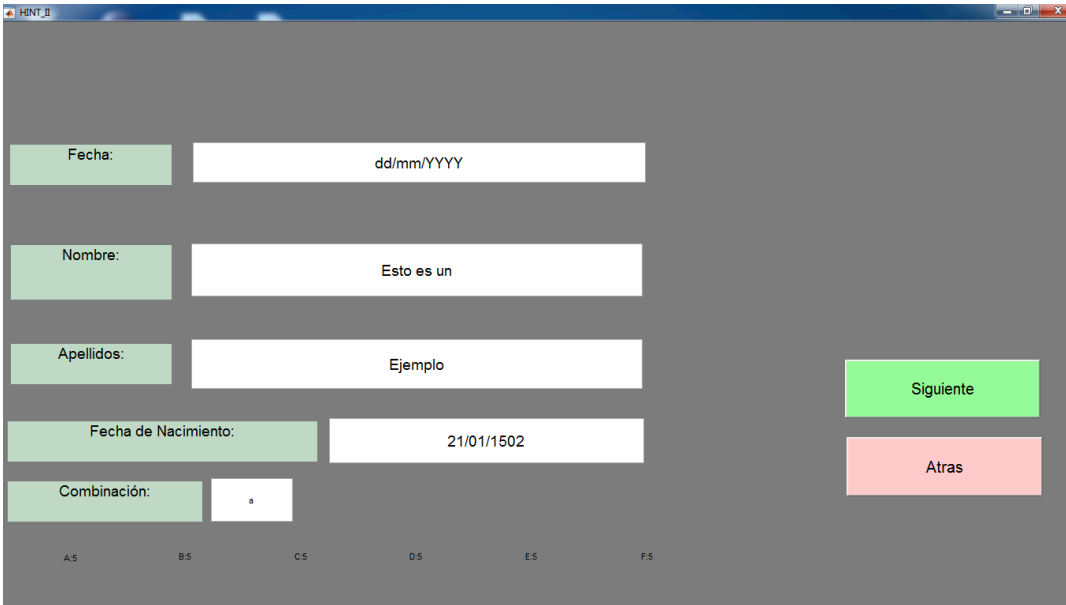
Para determinar si una respuesta es correcta o incorrecta, al encuestado escucha una frase, y este debe repetir lo que ha escuchado. El HINT especifica que para determinar que una frase es entendida correctamente, el encuestado debe acertar el 100% de sus palabras, pudiendo errar los artículos y en los 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.

La herramienta desarrollada tiene un entorno amigable. Cuenta con una página principal, donde se escoge la opción de empezar el test, o analizar frases, ruidos u otros test, además se ha añadido un mecanismo de control, para detectar outliers, (encuestados a los que descartaremos para realizar los test) al empezar los tests, para lo que se ha añadido un enlace en la página principal del programa que abre otra aplicación, esta permite realizar un simple test de audiometría entre las frecuencias 125 Hz y 8 KHz (decisión tomada tras el análisis de varias normas, como las UNE EN ISO 389 parte 1-2-3-4-5-6-7).. Además existe la posibilidad de guardar el grafico con el umbral auditivo para su análisis posterior.

Para detectar outliers se ha seguido el criterio de la OMS (Organización Mundial de la Salud) (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 (Figura 1), donde se recogen los datos de los usuarios a analizar: nombre, apellidos, fecha de nacimiento y fecha de realización del test.



The screenshot shows a software window titled 'JHINT\_II'. The interface is a data entry form with the following fields and values:

- Fecha:** dd/mm/YYYY
- Nombre:** Esto es un
- Apellidos:** Ejemplo
- Fecha de Nacimiento:** 21/01/1502
- Combinación:** a

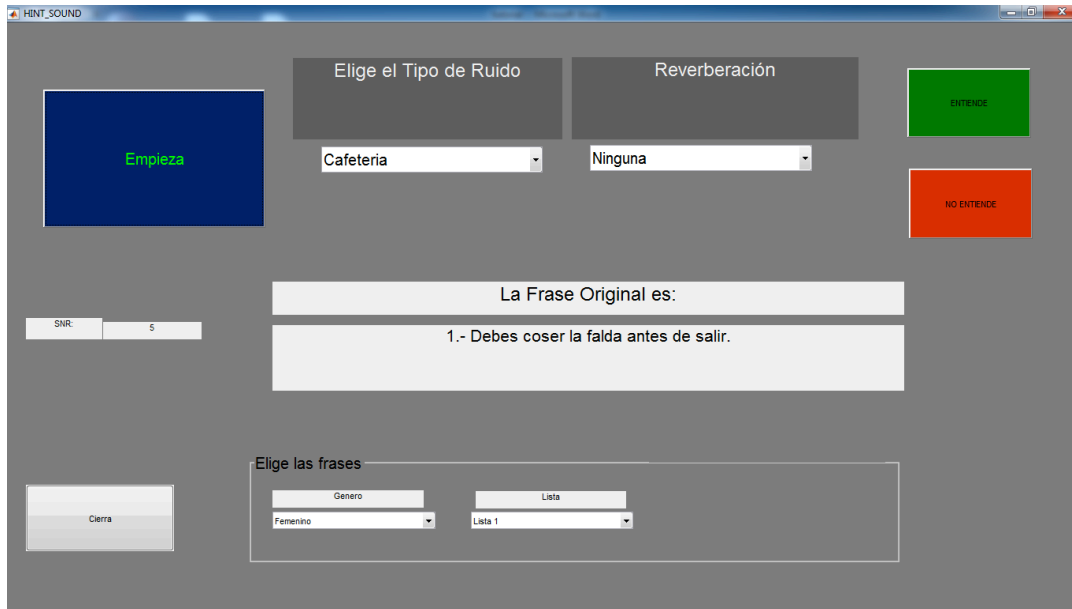
At the bottom of the form, there are six radio buttons labeled A.5, B.5, C.5, D.5, E.5, and F.5. On the right side, there are two buttons: a green 'Siguiente' button and a pink 'Atras' button.

Figura 1: Pagina 2 del JHINT

Se han utilizado 5 listas de 10 frases cada una, obtenidas del Sharvard Corpus [2] que serán reproducidas aleatoriamente por una voz masculina o por una voz femenina con un nivel de  $64\text{dB}_{\text{SPL}}$ .

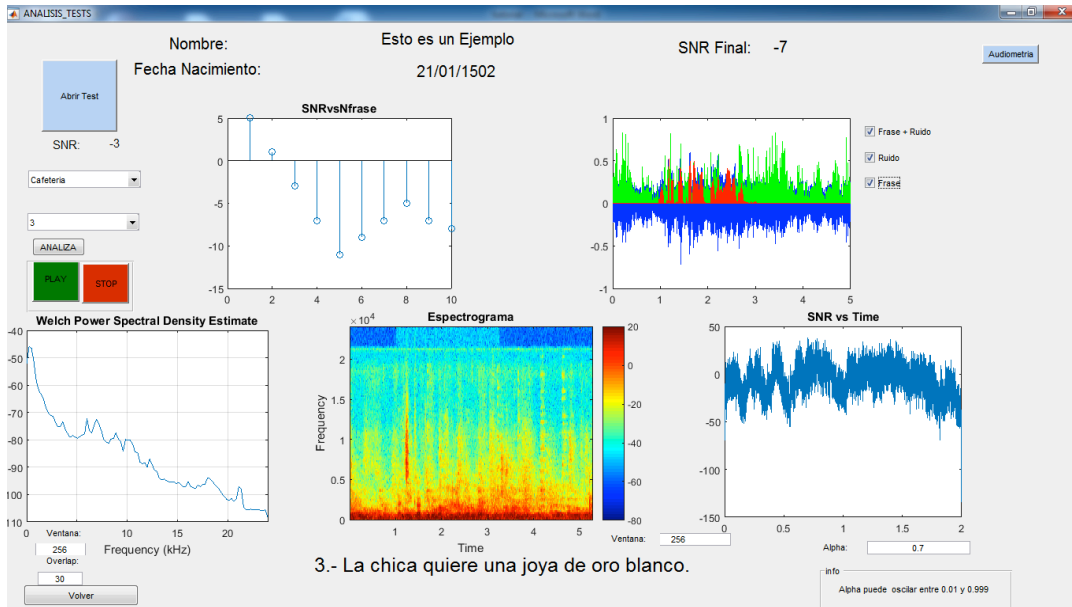
El procedimiento que se utiliza en la herramienta desarrollada es el mismo que para el HINT. Se ha decidido fijar una SNR inicial de 5 dB, debido a que es preferible empezar con una SNR favorable para el encuestado. Por lo tanto cuando empiece el test se reproducirá una frase con un ruido de fondo (de entre los ruidos considerados) y con una SNR de 5 dB. Seguidamente, la segunda, tercera y cuarta frases serán reproducidas incrementando o decrementando (dependiendo de si el usuario entiende correctamente o no dicha frase) su SNR 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  $\pm 1\text{dB}$ , consiguiendo un resultado final con mayor precisión.

En la figura 2 se visualiza el entorno realizado con *Matlab* de la aplicación JHINT, donde se realiza el procesado descrito anteriormente, el funcionamiento es el siguiente, primero se selecciona el tipo de ruido, seguidamente se selecciona la lista de frases y el genero de la voz que las va a reproducir, con los parámetros seleccionados, se pulsa el botón de **empieza**, este empezara a reproducir la frase, el encuestador debe juzgar leyendo la frase en pantalla si el encuestado acierta o no la misma. En caso de acertar se deberá pulsar **entiende**, y en caso de fallar, se pulsara **no entiende**.



**Figura 2:** Entorno grafico de la aplicación JHINT

Paralelamente, como vemos en la figura 3, se ha desarrollado una herramienta de análisis de los distintos tests con las siguientes prestaciones: tras abrir el test que queremos analizar, y seleccionar 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 puede visualizar la evolución de la potencia con el tiempo tanto para la frase como para el 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 3:** Herramienta de análisis JHINT

### 3. RESULTADOS OBTENIDOS DE LA APLICACIÓN JHINT

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. La aplicación JHINT se ha usado en una muestra formada por 30 individuos, con el fin de conseguir una SNR promedio para los distintos ruidos. Además para el ruido de cafetería se ha decidido procesar el mensaje para añadir una reverberación, teniendo para este ruido tres posibilidades: ruido de cafetería sin reverberación, con reverberación alta (TR60=0.7s) y con reverberación baja (TR60=0.2s). 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. Aleatoriamente se han asignados distintas combinaciones, con el fin de conseguir un mínimo de 10 tests realizados con los ruidos, sin tener que repetir el ruido de cafetería con las distintas reverberaciones a ningún usuario.

En la tabla 3 podemos observar los resultados obtenidos

**Resumen Estadístico para el valor de SNR buscado para cada ruido analizado.**

| Ruidos                       | Recuento | Promedio | Mediana | Moda | Varianza | Desviación Estándar |
|------------------------------|----------|----------|---------|------|----------|---------------------|
| '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              |

**Tabla 1 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 el valor de SNR buscado para cada ruido analizado.**

| 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              |

Tabla 2 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.

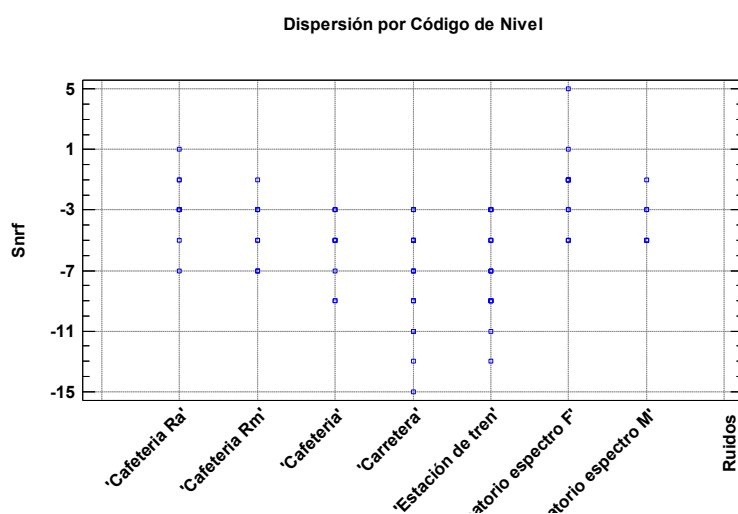


Figura 4: Dispersión por ruido

En la figura 4 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 5, 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.

Gráfico Caja y Bigotes

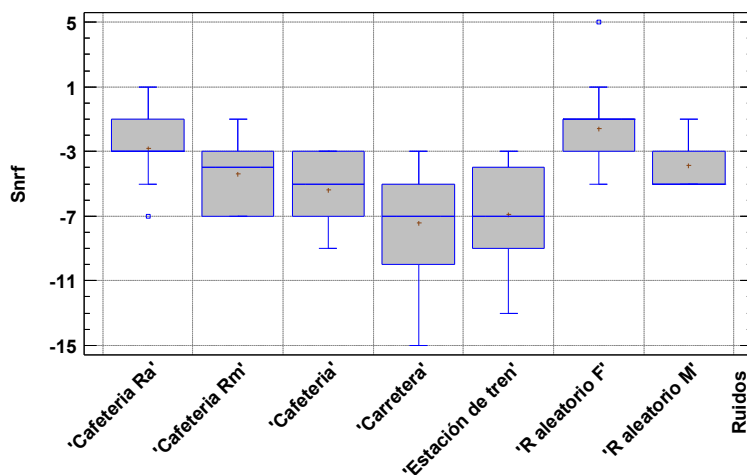


Figura 5: Grafico de cajas y bigotes

#### 4. CONCLUSIONES, APLICACIONES Y POSIBLES MEJORAS

Tras haber estudiado diversos procedimientos para calcular la inteligibilidad mediante percepción del habla, se ha decidido realizar una herramienta que permite calcular el nivel de SNR para conseguir el 50 % de inteligibilidad. Esta herramienta permite evaluar la inteligibilidad no sólo 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 (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 límites tolerables.

La herramienta descrita podría ser utilizada con otros ruidos cotidianos que fueran de interés.

Por otra parte una posible aplicación de 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, esta herramienta permitirá configurar adecuadamente los audífonos, ya que puede emplearse en los usuarios de estos dispositivos (por ejemplo, podemos encontrar la mejor configuración de un audífono realizando el test a la persona que lo utiliza, con distintas configuraciones y encontrando cual es la que proporciona una mejor inteligibilidad o una SNR más baja). Por tanto, a través del análisis de los resultados para las diferentes configuraciones de los audífonos obtenidos mediante esta aplicación se podrían evaluar las mejoras en el procesado de dichos audífonos.

Otra de posibilidades que ofrece la aplicación sería tratar de emular el test que realiza Lena L.N. Wong 2012, [5] 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 estudios 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 [5] 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.

## **5. AGRADECIMIENTOS**

Este trabajo ha sido financiado por la Unión Europea FEDER y el Gobierno español a través del proyecto TEC2012-38142-C04, y la Generalitat Valenciana a través del proyecto PROMETEOII/2014/003 y la ayuda FPA/2015/032.

## **6. REFERENCIAS BIBLIOGRAFICAS**

[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] 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

[3] ANSI. (1997) *S3.5 American National Standard Methods for Calculation of the Speech Intelligibility Index*. (Acoustical Society of America, New York)

[4] Bamford-Kowal-Bank (BKB) (Bench y Bamford, 1979)

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