



Interdisciplinaria

ISSN: 0325-8203

ISSN: 1668-7027

[interdisciplinaria-ciipme@conicet.gov.ar](mailto:interdisciplinaria-ciipme@conicet.gov.ar)

Centro Interamericano de Investigaciones Psicológicas y Ciencias Afines

Argentina

Maggi, Ana Luz; Evin, Diego; Hinalaf, María  
Evaluación de la discriminación del habla en ruido:  
Equivalencia de listas BEPPA en sujetos con audición normal  
Interdisciplinaria, vol. 41, núm. 2, 2024, Mayo-Diciembre, pp. 23-24  
Centro Interamericano de Investigaciones Psicológicas y Ciencias Afines  
Buenos Aires, Argentina

Disponible en: <https://www.redalyc.org/articulo.oa?id=18081840023>

- ▶ [Cómo citar el artículo](#)
- ▶ [Número completo](#)
- ▶ [Más información del artículo](#)
- ▶ [Página de la revista en redalyc.org](#)

The logo for Redalyc.org, featuring the text 'redalyc.org' in a lowercase, sans-serif font. The 'red' part is in red, and the 'alyc.org' part is in black. A small red diamond shape is positioned above the 'y'.

Sistema de Información Científica Redalyc  
Red de revistas científicas de Acceso Abierto diamante  
Infraestructura abierta no comercial propiedad de la academia

**Evaluación de la discriminación del habla en ruido: Equivalencia de listas BEPPA  
en sujetos con audición normal**

**Assessment of speech discrimination in noise: Equivalence of BEPPA lists in  
normal hearing subjects**

Ana Luz Maggi<sup>1</sup>, Diego Evin<sup>2</sup> y María Hinalaf<sup>3</sup>

<sup>1</sup>Universidad Nacional de Córdoba, Argentina. <https://orcid.org/0000-0001-6683-7493>.

E-mail: [ana.luz.maggi@unc.edu.ar](mailto:ana.luz.maggi@unc.edu.ar)

<sup>2</sup>Universidad Nacional de Entre Ríos, Argentina.

<https://orcid.org/0000-0002-4903-7897>. E-mail: [diegoevin@gmail.com](mailto:diegoevin@gmail.com)

<sup>3</sup>Universidad Nacional de Córdoba, Argentina. <https://orcid.org/0000-0003-1172-4306>.

E-mail: [maria.hinalaf@unc.edu.ar](mailto:maria.hinalaf@unc.edu.ar)

La presente investigación se realizó con un subsidio Consolidar de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba y una Beca Interna Doctoral del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Centro de Investigación y Transferencia en Acústica (CINTRA, CONICET-UTN).

Córdoba, Argentina.

**Resumen**

La dificultad para discriminar el habla inmersa en ruido de fondo provoca molestias, pérdidas funcionales en actividades cotidianas y es un indicador prematuro de deterioros auditivos. Su evaluación cobra cada vez más relevancia. Sin embargo, actualmente en Argentina, son escasos los protocolos estandarizados para su empleo en la rutina audiológica. Este trabajo tiene por objetivo estudiar la equivalencia de las diferentes listas de la batería BEPPA para su adopción como instrumento de evaluación

audiológica. Se realizó un estudio descriptivo correlacional transversal de discriminación del habla en 32 jóvenes con umbrales de audición normales, para lo cual se utilizaron listas de palabras monosílabas, bisílabas, con secuencias consecutivas de vocales o transiciones vocálicas, y oraciones de un grupo entonativo de la batería BEPPA. Esos estímulos, pronunciados por una locutora profesional, se presentaron a los participantes de forma binaural diótica sin ruido y con ruido de características espectrales similares a las del habla a una relación señal ruido (RSR) de 5dB, 0dB y -5dB, para solicitar a los participantes que repitieran luego de cada estímulo lo que escucharan en voz alta. A fin de considerar el factor de frecuencia de uso en la activación léxica, se registró la frecuencia de cada palabra utilizando como referencia dos corpus lingüísticos del español contemporáneo. Los resultados mostraron diferencias estadísticamente significativas en cantidad de errores al contrastar las condiciones RSR 5dB y 0dB; y entre condiciones 0dB y -5dB, lo que indica que resultan suficientemente diferentes como contextos de evaluación. El contraste de frecuencias léxicas entre las listas de palabras de cada categoría no mostró diferencias estadísticamente significativas, por lo que se podría descartar el efecto en las respuestas de este elemento como factor de confusión. A partir de las diferencias observadas en los errores cometidos, se pudo concluir que, a excepción de dos listas de palabras monosílabas, las demás muestran dificultades de reconocimiento equivalentes y, por lo tanto, resultan útiles para la evaluación audiológica.

*Palabras clave:* audición, pruebas auditivas, discriminación del habla en ruido, audiometría del habla, adulto joven

### **Abstract**

The difficulty in discriminating speech immersed in background noise is a condition that, besides causing discomfort and functional loss in daily activities, is an early

indicator of hearing impairment. The evaluation of this hearing function has gained relevance in the field of audiology in recent years, however, there are currently few standardized protocols for its use as a routine audiological screening in Argentina. The purpose of this work is to study the equivalence of the different lists of the BEPPA battery with the aim of adopting it as an assessment tool in the audiology routine. This research was carried out at CINTRA – UE CONICET UTN. A cross-sectional descriptive correlational study of speech discrimination was conducted on 32 young people with normal hearing thresholds UNC and UTN students during the years 2019 and 2020. Normal tonal thresholds ( $\leq 21$  dB) within the conventional range (250–8000) Hz, hearing rest for at least 8 hours and signed informed consent were established as inclusion criteria. The research was approved by the Institutional Committee for Ethics in Health Research of the Hospital Nacional de Clínicas. Lists of monosyllabic and bisyllabic words, words with consecutive vowel sequences or “vocalic transitions”, and sentences with a single intonational group from the BEPPA battery were used. These stimuli were recorded in a CINTRA soundproof chamber, by a professional female speaker and masked using noise with a speech-like spectrum. The signals were reproduced by a digital audiometer. were presented to participants in a binaural dichotic manner in quiet and masked with noise of similar spectral characteristics to those of speech at a signal-to-noise ratio (SNR) of 5dB, 0dB and -5dB, asking them to repeat after each stimulus what they heard aloud. To consider the frequency of use factor on lexical activation, the frequency of each word was recorded using two linguistic corpora of contemporary Spanish as a reference. The results showed statistically significant differences in the number of errors when contrasting the RSR conditions 5dB and 0dB; and between 0dB and -5dB conditions, indicating that they are different enough as evaluation contexts. The contrast of lexical frequencies between the word lists of each

category did not show significant differences, so its effect as a confounding factor in the responses obtained could be ruled out. From the observed differences in errors between lists for each category, we can conclude that except for two monosyllabic word lists, the others exhibit equivalent recognition difficulties and are therefore useful for audiological assessment.

*Keywords:* hearing, hearing tests, speech perception in noise, speech audiometry, young adult

### **Introducción**

La capacidad de discriminar el habla cuando está inmersa en ruido depende tanto de factores perceptuales auditivos (*bottom-up*) como de aspectos cognitivos (*top-down*) (Rönnerberg et al., 2010). La dificultad para discriminar el habla en ruido es un síntoma frecuente que puede indicar prematuramente alguna discapacidad auditiva, además de provocar molestias y pérdidas funcionales en las actividades cotidianas. La Asociación Americana de Audiología (AAA) y la Asociación Americana del Habla, Lenguaje y Audición (ASHA) han sugerido, en sus guías de buenas prácticas, que las pruebas de discriminación del habla en ruido (DHR) formen parte de una evaluación audiológica integral. Aun así, menos del 15 % de especialistas en audiología utilizan estas pruebas de forma rutinaria (Beck et al., 2018). Actualmente la DHR es abordada principalmente en personas que poseen prótesis auditivas, para poder calibrar los equipos de la forma más favorable y realizar un seguimiento. En Argentina, la rutina audiológica básica sí incluye la evaluación de discriminación del habla pero sin ruido, a través de la logaudiometría convencional. La discriminación del habla sin ruido y la DHR no presentan variaciones proporcionales: dos individuos pueden tener la misma pérdida auditiva en relación a sus umbrales tonales audiométricos, el mismo desempeño en discriminación del habla sin ruido y, sin embargo, sus desempeños en la DHR pueden

ser sustancialmente diferentes (Taylor, 2003; Wilson y McArdle, 2008). Es por ello que resulta importante poder incorporar en Argentina la evaluación de esta función auditiva en pacientes sin prótesis auditivas como parte de la rutina audiológica.

Para la lengua inglesa, la discriminación del habla se ha investigado extensamente desde la década de 1940 y se han registrado y estandarizado los materiales para su evaluación (Carlo et al., 2020). En relación a la discriminación del habla, específicamente en presencia de ruido competitivo, existen diversas pruebas en el habla inglesa. Entre las más reconocidas se encuentran Speech Perception in Noise (SPIN) (Kalikow y Stevens, 1977), Connected Speech Test (CST) (Cox et al., 1987), Speech in Noise Test (SIN) (Killion y Vilchur 1993), Hearing in Noise Test (HINT) (Nilsson et al., 1994), Quick Speech in Noise Test (QuickSIN) (Killion et al., 2004), Speech Recognition in Noise Test (SPRINT) (Cord et al., 1992), Words In Noise Test (WIN) (Wilson, 2003).

Estas pruebas difieren entre sí respecto al tipo de estímulo a discriminar (*target*), el ruido competitivo, el método de ejecución y la forma de calcular los resultados (Tabla 1). Algunas de las pruebas utilizan estímulos que son más representativos de la DHR en la vida diaria como el HINT que utiliza oraciones; en otros casos los estímulos resultan menos ecológicos, como es el caso del WIN que utiliza monosílabos aislados. Con relación al ruido competitivo, la mayoría de las pruebas utilizan ruido conversacional multihablante, que es la combinación aditiva de las emisiones simultáneas de varios hablantes. La ventaja de este tipo de ruido es que resulta más natural, es decir, similar a las situaciones cotidianas. Sin embargo, al ser variable, fluctúa y puede presentar picos de enmascaramiento. Respecto a la forma en que se presentan los estímulos, hay clásicamente dos estrategias: la de Relación Señal Ruido (RSR) fija y la adaptativa. En las pruebas de RSR fija se usan niveles predeterminados de RSR para un grupo de estímulos y se obtiene el porcentaje o número bruto de repeticiones correctas e

incorrectas. Entre las pruebas que hacen uso de esta estrategia se encuentran CST -Connected Speech Test- y SPIN –Speech Perception in Noise. En el caso de la RSR adaptativa, la relación entre los niveles de habla y ruido aumenta o disminuye dependiendo del rendimiento del paciente. En general, se evalúan distintos niveles de RSR y luego se calcula o infiere a que RSR el sujeto reconoce correctamente el 50 % de los estímulos, a lo que se denomina umbral RSR. Las pruebas HINT y QuickSIN adoptan esta estrategia.

**Tabla 1.**

***Principales características de pruebas que evalúan la discriminación del habla en ruido en inglés***

<b>Prueba</b>	<b>Referencia</b>	<b>Target</b>	<b>Ruido competitivo</b>	<b>Método</b>	<b>Resultado</b>
<b>Speech Perception in Noise (SPIN)</b>	Kalikow, Stevens y Elliot (1977)	Última palabra (monosílabos) de oraciones. Se utilizan oraciones de alta predictibilidad y baja predictibilidad	Conversacional 12 hablantes	RSR fija en 0 dB. Habla y ruido a 50 dB HL	Porcentaje de palabras de alta predictibilidad y baja predictibilidad repetidas correctamente
<b>Connected Speech Test (CST)</b>	Cox, Alexander y Gilmore (1987)	Palabras claves de pasajes de habla conectada producida conversacionalmente	Conversacional multihablante	RSR fija	Promedio del porcentaje de repeticiones correctas en 4 listas.
<b>Speech Recognition in Noise Test (SPRINT)</b>	Cord, Walden y Atack (1992)	Palabras monosilábicas aisladas	Conversacional 6 hablantes	RSR fija en 9 dB. Habla a 50 dB HL y ruido a 41 dB HL	Número bruto de palabras repetidas correctamente
<b>Speech in Noise Test (SIN)</b>	Killion y Vilchur (1993)	Cinco palabras claves por cada oración	Conversacional de 4 hablantes	RSR adaptativa en condición de RSR 15, 10, 5 y 0 dB. Dos posibilidades de evaluación habla fija a 80 dB SPL y	Umbral del 50 % de reconocimiento del habla.

				habla fija a 50 dB SPL	
<b>Hearing in Noise Test (HINT)</b>	Nilsson, Sali y Sullivan (1994)	Todas las palabras de cada oración	Ruido continuo con características espectrales similares al habla (SSN)	RSR adaptativa, con ruido fijo a 65 dB(A) variando habla con saltos de 4 dB y luego 2 dB	Umbral del 50 % de reconocimiento del habla
<b>Words in noise test (WIN)</b>	Wilson (2003)	Palabras monosilábicas aisladas	Conversacional de 6 hablantes	RSR adaptativa de 24 dB RSR a 0 dB RSR con saltos de 2 dB. Ruido fijo a 60 dB HL.	Umbral del 50 % de reconocimiento del habla
<b>QuickSIN (Derivación del SIN)</b>	Killion, Niquette y Gudmundsen (2004)	Cinco palabras claves por cada oración	Conversacional de 4 hablantes	RSR adaptativa en: 25, 20, 15, 10, 5 y 0 dB. Se sugiere ruido fijo a 70 dB HL.	Valor de pérdida RSR (SNRloss). Se realiza una resta, entre el valor normativo y el obtenido por el paciente.

Para el caso del habla hispana son pocas las pruebas propuestas para evaluar la discriminación del habla en ruido. En 2002 se desarrolló una versión en español de la prueba HINT (Soli et al., 2002); sin embargo, esta versión requiere contar con hardware y software específicos, lo que limita su adopción en la clínica. Otra experiencia en este ámbito fue la realizada en España por el Programa Infantil Phonak (PIP) y la Universidad Nacional de Educación a Distancia (UNED) (Maggio et al., 2016), (Marrero-Aguilar, 2015), quienes crearon una batería de frases para logaudiometría infantil en ruido para controlar las variables de equilibrio fónico, frecuencia y familiaridad del léxico, la estructura sintáctica y el patrón entonativo. Para que ese material pueda ser utilizado en otro país, necesitarían de una adaptación cultural. En Argentina, Gurlekian et al. (2008) desarrollaron una prueba rápida de discriminación del habla en ruido para niños de entre 6 a 12 años, diseñada con el objetivo de aplicarla de forma sencilla en establecimientos escolares. La prueba utiliza como enmascarador

un ruido modulado con forma de habla (Modulated Speech-Shaped Noise, MSSN) y como *target*, oraciones con una estructura sintáctica sencilla del tipo sujeto-verbo-objeto. Las palabras utilizadas fueron tomadas de manuales de primer grado. Las oraciones se construyeron de forma que no superasen 7-8 palabras. Luego, a partir de un número considerable de oraciones, se obtuvieron grupos de oraciones fonéticamente balanceadas de acuerdo con la descripción de los fonemas y alófonos más frecuentes de Argentina.

Por su parte, Aronson et al. (2007) desarrollaron la Batería para Pacientes con Prótesis Auditiva (BEPPA). Este corpus, diseñado en base al español rioplatense, permite la evaluación de la discriminación del habla, con y sin ruido competitivo, en adultos equipados con audífonos o implantes cocleares. El tipo de material que presenta es diverso: consonantes en contexto vocálico, palabras monosilábicas, palabras bisilábicas, palabras con secuencias vocálicas consecutivas (transiciones vocálicas), oraciones de un grupo entonativo y otras de múltiples grupos entonativos. Las listas se encuentran fonéticamente balanceadas. Se realizaron grabaciones con voces de tres hablantes femeninos y tres hablantes masculinos, nativos de Argentina. Se generaron pistas a -5, 0, 5, 10 y 15 dB de RSR, para las que se utilizó ruido blanco y ruido rosa. En esta batería, el mecanismo de selección de los ítems a evaluar y la metodología de evaluación quedan sujetos al criterio profesional, en el que su mayor beneficio reside en su versatilidad. En la presente investigación se adoptará esa batería, ya que es la que presenta mayor diversidad de estímulos, lo que resulta conveniente para los objetivos de investigaciones actuales y futuros.

Una característica a considerar cuando se evalúa la discriminación del habla es la frecuencia léxica o frecuencia de ocurrencia de la palabra del material a utilizar. Esta es definida como la tasa de uso de una palabra en una lengua. Es probable que las palabras

más frecuentes sean reconocidas con mayor facilidad por los sujetos (Bargetto y Riffo Ocares, 2019; Luce y Pisoni 1998). Es decir que, a la hora de evaluar la discriminación del habla, las palabras muy frecuentes tienen una ventaja permanente en el proceso de activación léxica. Este trabajo tiene por objetivo estudiar la equivalencia de las diferentes listas de la batería BEPPA, un primer paso para validar su adopción en pruebas de DHR en personas con audición dentro de los parámetros considerados normales. Para ello, se realizó una primera prueba piloto en sujetos normooyentes, en la que se propuso: a) comparar la cantidad de errores entre las distintas condiciones de RSR evaluadas; b) analizar y comparar la cantidad de errores de cada lista; c) analizar la relación entre la cantidad de errores de cada palabra y la frecuencia léxica de las palabras registrada en un corpus de habla del español, y d) comparar las distintas listas según su frecuencia léxica.

Los objetivos de este estudio contemplan a futuro con esta batería, por un lado, poder indagar la relación del sistema eferente medial y la DHR y, por el otro, poder analizar las incipientes dificultades en DHR, cuando todavía no se ha manifestado una hipoacusia.

### **Método**

Se llevó a cabo un estudio descriptivo correlacional transversal sobre una muestra piloto de 32 estudiantes de grado de la Universidad Nacional de Córdoba (UNC) y la Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN-FRC) durante los años 2019 y 2020. Los participantes fueron estudiantes que se presentaron como voluntarios antes una convocatoria en el ámbito universitario. Como criterio de inclusión se estableció que los participantes debían tener umbrales tonales considerados normales ( $\leq 21$  dB) dentro del rango convencional (250–8000) Hz; haber realizado reposo auditivo de al menos 8 horas, y firmar el consentimiento informado.

La evaluación auditiva se realizó en una cabina audiométrica que se ajusta a la norma IRAM 4026:1986 (Gaetán et al., 2019).

El trabajo de campo fue evaluado y aprobado por el Comité Institucional de Ética de las Investigaciones en Salud (CIEIS) del Hospital Nacional de Clínicas. A cada participante se le informó de forma oral y escrita los procedimientos que se llevarían a cabo, se les explicó que no participar no traería ningún perjuicio y que podían retirarse de las pruebas en cualquier momento que lo desearan. Luego de esto, se les pidió que firmaran un consentimiento informado.

### **Evaluación de la discriminación del habla**

Se utilizó el corpus de la batería BEPPA, diseñada por Aronson et al., (2007). Las listas fueron diseñadas y ajustadas por los autores para que sean fonéticamente balanceadas. En el presente estudio se emplearon las listas de palabras monosílabas, bisílabas, palabras con transiciones vocálicas (TV) y oraciones de un grupo entonativo (OUGE) de la batería BEPPA (Tabla 2).

**Tabla 2.*****Material de la batería BEPA utilizado en la investigación***

<b>Tipo de estímulo verbal</b>	<b>Número de Listas</b>	<b>Elementos por Lista</b>	<b>Ejemplos</b>
<b>Monosílabas</b>	6	15 pal.	<i>mar, sed, paz, ven, gran, red, ...</i>
<b>Bisílabas</b>	6	10 pal.	<i>asma, alma, asta, orca, isla, ...</i>
<b>TV</b>	6	19 pal.	<i>aéreo, aire, ahora, auto, ...</i>
<b>OUGE</b>	6	10 orac.	<i>Sesenta gramos, ¿Te gustó?...</i>

Dicho material fue presentado a cada sujeto bajo las siguientes condiciones: sin ruido de fondo y con ruido a una RSR de 5dB, 0dB y -5dB. En cada condición, se brindó una lista distinta elegida a través de un muestreo aleatorio sistemático para evitar que el sujeto las memorice. A cada participante se le brindó cuatro de las seis listas disponibles y se procuró que todas las listas fueran empleadas en todas las condiciones para luego poder compararlas. Los estímulos se emitieron de forma binaural diótica.

Simultáneamente a la presentación de los estímulos, se solicitó a los participantes que inmediatamente oído cada estímulo lo repitieran en voz alta. Estas respuestas fueron grabadas acústicamente en una grabadora de voz Zoom modelo H4n, y se utilizaron posteriormente para evaluar los aciertos en la discriminación de los estímulos verbales. Se consideraron errores aquellos casos en que las repeticiones en voz alta de las palabras contuviesen omisiones, sustituciones o inserciones de uno o más fonemas. En el caso de las OUGE fue imprescindible calcular el porcentaje de palabras con error, ya que las oraciones y las listas tenían algunas leves variaciones en cuanto a la cantidad de palabras que las componían. Las señales fueron reproducidas a través de audiómetro digital Madsen, modelo Orbiter 922 y auriculares Sennheiser, modelo HDA 200.

***Grabación y procesamiento digital***

Las grabaciones de las listas BEPPA fueron realizadas por una locutora profesional en una cámara insonorizada del Centro de Investigación y Transferencia en Acústica (CINTRA).

Se utilizó un micrófono de vincha Beyerdynamic TG H75HC, de tipo electret, omnidireccional. Para la adquisición y digitalización de las señales, se usó la placa Focusrite Scarlett 18i8 2da Gen, en modo monoaural, 44100 Hz de frecuencia de muestreo y resolución de 24 bits por muestra. Para completar la cadena de adquisición se empleó una notebook Lenovo Ideapad 720, y el software Audition V.

A diferencia de la propuesta original, en este trabajo se decidió emplear, como señal enmascarante, ruido continuo con características espectrales similares al habla (SSN) en lugar del ruido blanco y ruido rosa con los que originalmente se concibió la batería.

Como se mencionó, originalmente esta batería se propuso para evaluar pacientes con implantes cocleares. En el caso de esta investigación, se adoptó la decisión de cambiar el tipo de ruido por la necesidad de generar situaciones de mayor competitividad, acorde a su aplicación en sujetos con audición normal. Entre las alternativas de ruidos más competitivos, Wilson (2007) comparó el rendimiento de DHR al utilizar ruido conversacional multihablante y ruido SSN, y encontraron que este último era más efectivo como enmascarador.

El SSN se generó a partir de las señales limpias de la batería y el siguiente procedimiento: en primer lugar, se obtuvo la transformada de Fourier de la señal conjunta de todos los archivos de la batería para la locutora, luego se aplicó un preprocesamiento para hacer que las fases de las componentes espectrales resulten aleatorias y, finalmente, se aplicó la transformada inversa para llevarla a dominio temporal. Para el proceso de adición de ruido a las señales de habla limpias en las diferentes RSR se utilizó la función “v\_addnoise” del Toolbox para procesamiento de

voz: VoiceBOX (Brookes, 2009), que calcula el nivel de la señal vocal al usar una medida de nivel activo del habla de acuerdo con la Recomendación UIT-T P.56 (UIT-T, 1993).

Para cada lista de la batería se diseñó una versión sin ruido y versiones con ruido a una RSR de 5, 0 y -5 dB, en el que el nivel de habla se mantuvo fijo y se realizaron variaciones en el nivel de ruido. En la BEPPA original se generaron pistas a -5, 0, 5, 10 y 15 dB de RSR; sin embargo, para este trabajo se decidió evaluar solamente las pistas sin ruido y con RSR a 5, 0 y -5 dB. El valor mínimo fue seleccionado y se tuvieron en cuenta las investigaciones previas (Armstrong et al., 2020; Koole et al., 2016) que mencionan que entre 0 y -5 dB de RSR se encuentra el punto de corte que permite detectar la audición anormal en pruebas de percepción del habla. Por su parte, el valor máximo se estableció al considerar que al evaluar jóvenes con umbrales tonales dentro de los parámetros normales, resultaba esperable que resultados de las pruebas con RSR a 10 y 15 dB exhibiesen un efecto techo (Holder et al., 2018), es decir, que no se observaran errores de inteligibilidad. De esta manera, se decidió evitar esas pistas que brindarían resultados con poca o nula variabilidad y priorizar la variedad de tipos de estímulos verbales a evaluar. En este sentido, es importante señalar que para su aplicación en audiología clínica, la extensión de la evaluación es un factor relevante de uso.

Para compilar los grupos de listas, entre cada estímulo se dejó un tiempo de silencio o ruido, según cuál fuera la condición de la prueba para que los participantes pudieran pronunciar el estímulo oído. Para evitar que el período de respuesta se superponga con el de estimulación, el tiempo se definió como una vez y media la duración del componente más largo de cada categoría.

### ***Calibración de la cadena de ensayos***

La cadena de ensayos –compuesta por una PC, un audiómetro y auriculares–, se calibró para obtener niveles de presión sonora de 60 dB. Es decir, la salida de señal del habla se estableció a 60 dB. Para ello, se utilizó un oído artificial marca Brüel y Kjaer 4153 con acoplador IEC, micrófono de la misma marca modelo 4134 y un medidor de nivel sonoro de la misma marca y modelo 2250. Se fijaron los niveles de audio en el reproductor de la PC (Windows Media) en 100 % y de la salida de la placa de sonido (altavoces para el sistema operativo Windows 10) en 60 %, valores que se debería volver a verificar para cada ensayo. Finalmente, se utilizó el seteo de HL del audiómetro para ajustar iguales niveles de presión sonora en los dos canales, izquierdo y derecho. Los auriculares marca Sennheiser modelo HDA 200 se ajustaron al oído artificial de manera tal que la presión en almohadillas fuera la correspondiente a la fuerza de ajuste de vincha recomendada por Norma que es de  $10 \pm 1$  N.

### **Audiometría**

Se evaluó el rango convencional (250–8000) Hz para determinar que los umbrales auditivos estuvieran dentro de los parámetros considerados normales ( $\leq 21$  dB). Se evaluaron ambos oídos por vía aérea, mediante tonos puros y variaciones de intensidad de 3 dB HL. Se utilizó un audiómetro Madsen Orbiter 922 con auriculares supraaurales Sennheiser modelo HDA 200. Antes de realizar la audiometría se realizó una visualización del conducto auditivo externo para corroborar que no estuviera ocupado, para lo que se empleó un otoscopio modelo Heine Beta 100.

### **Frecuencia léxica**

Se registró la frecuencia de ocurrencia de cada palabra de las listas de bisílabos, monosílabos y TV. Para esto se utilizaron dos grandes corpus lingüísticos. Por un lado, el Corpus del Español NOW (Davies, 2018), conformado por alrededor de 7 billones de palabras, tomados de periódicos y revistas *online* entre los años 2012 y 2019, de 27

países de habla hispana, entre los que se encuentra Argentina. Por otro lado, se utilizó el Corpus del Español del Siglo XXI (CORPES XXI) en su versión beta 0.99 (Real Academia española), que cuenta con más de 316 000 documentos, que suman algo más de 333 millones de formas ortográficas, procedentes de textos escritos y de transcripciones de textos orales provenientes de España, América, Filipinas y Guinea Ecuatorial, desde el año 2001. La ventaja de este corpus es que permite seleccionar un subcorpus específicamente de Argentina y, en este caso, ajustar la comparación de frecuencias léxicas con evidencias más representativas de la región.

### **Análisis estadísticos**

El análisis estadístico se realizó mediante el software Infostat versión 2020 (Di Rienzo et al., 2020). Se utilizó el Test Shapiro-Wilks para conocer si las variables presentaban una distribución normal. Para comparar el nivel de reconocimiento de cada condición de RSR se aplicó el Test Wilcoxon, con un nivel de significación de  $p < .05$ . Se realizaron gráficos de media y desvío estándar (DE) de porcentajes de aciertos según cada lista. Se utilizó el test Kruskal-Wallis para comparar la cantidad de errores de cada lista y se consideraron todas las condiciones de RSR en conjunto y para comparar la frecuencia léxica entre las listas. Se realizaron tablas en las que se analizaron las palabras y oraciones con más de tres errores y se tuvieron en cuenta todas las condiciones de RSR. Para estudiar la correlación entre los errores de palabras y la frecuencia léxica se analizó el coeficiente de Pearson con un nivel de significación de  $p < .05$ .

### **Resultados**

La muestra quedó conformada por 32 jóvenes normooyentes con edades comprendidas entre 21 y 30 años, con una media de 21.38 (DE = 2.26). De estos participantes, 21 se identificaban con el género femenino (66 %) y 11 con el género masculino (34 %).

En la Tabla 3 se presentan los análisis de cantidad de errores según condición de RSR para los distintos tipos de estímulos verbales. Tanto en monosílabos, bisílabos, TV y OUGE se puede observar que los errores en la discriminación aumentan a medida que el ruido se vuelve más competitivo. También se puede apreciar que, en general, a menor contexto del estímulo (menor duración), mayor deterioro en la tasa de reconocimiento al aumentar el ruido de fondo. Una excepción a ese comportamiento se observa en el caso de RSR -5 en la que la tasa de error de bisílabos supera a monosílabos.

**Tabla 3.**

*Errores en la discriminación de monosílabos, bisílabos, transiciones vocálicas (TV) y oraciones de un grupo entonativo (OUGE) según condición de RSR y comparación de las distintas condiciones*

Condición RSR	n	Porcentaje de errores	Media de errores	Desvío estándar	p valor
<b>MONOSÍLABOS</b>					
Sin ruido	32	4.58	.69	1.18	.0168* (Vs RSR 5)
RSR 5	32	8.12	1.22	1.18	< .0001* (Vs RSR 0)
RSR 0	32	21.87	3.28	2.37	< .0001* (Vs RSR -5)
RSR -5	32	55.41	8.31	2.22	-
<b>BISÍLABOS</b>					
Sin ruido	32	1.25	0.13	0.34	.2608 (Vs RSR 5)
RSR 5	32	3.75	0.38	0.49	< .0001* (Vs RSR 0)
RSR 0	32	17.81	1.78	1.04	< .0001* (Vs RSR -5)
RSR -5	32	63.43	6.34	2.19	-
<b>TRANSICIONES VOCÁLICAS</b>					
Sin ruido	32	0.98	0.19	0.47	.0628 (Vs RSR 5)
RSR 5	32	2.46	0.47	0.67	< .0001* (Vs RSR 0)
RSR 0	32	11.84	2.25	1.41	< .0001* (Vs RSR -5)
RSR -5	32	46.21	8.78	2.92	-
<b>ORACIONES UN GRUPO ENTONATIVO</b>					

<b>Sin ruido</b>	32	.00	-	.00	.0108* (Vs RSR 5)
<b>RSR 5</b>	32	1.31	-	2.34	.0112* (Vs RSR 0)
<b>RSR 0</b>	32	3.33	-	4.46	< .0001* (Vs RSR -5)
<b>RSR -5</b>	32	30.05	-	14.12	-

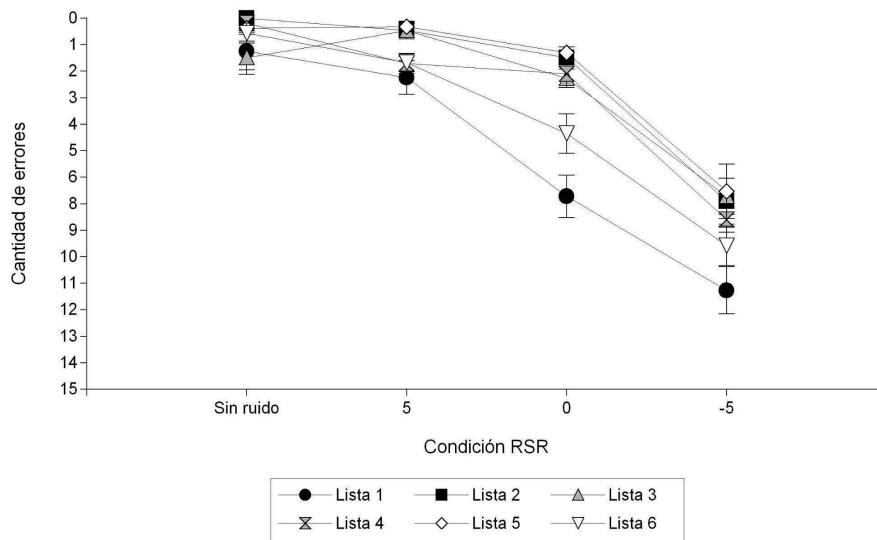
Test de Wilcoxon para muestras apareadas, se compararon los errores en la discriminación de bisílabos de cada condición con la siguiente condición más competitiva: Sin ruido vs RSR 5; RSR 5 vs 0; RSR 0 vs -5.

Con el propósito de conocer la significancia estadística de estas diferencias, se realizó el test Wilcoxon para muestras apareadas. En los monosílabos se observaron diferencias estadísticamente significativas ( $p < .05$ ) para todas las comparaciones de condiciones RSR. En los bisílabos no se registraron diferencias significativas al comparar la condición sin ruido respecto a la RSR 5. Cuando se comparó RSR 5 con RSR 0 y RSR 0 con RSR -5 se obtuvieron diferencias estadísticamente significativas ( $p < .05$ ). En las TV no se observaron diferencias estadísticamente significativas ( $p > .05$ ) entre la condición sin ruido y RSR 5, pero sí diferencias entre RSR 5 y 0dB, y entre RSR 0 y -5dB. Al analizar las OUGE se encontraron diferencias estadísticamente significativas ( $p < .05$ ) para todos los pares de condiciones analizados.

Al comparar los porcentajes de errores en RSR -5 dB de los distintos estímulos verbales, se pudo observar que las OUGE tuvieron el menor porcentaje de errores con 30.05 %, seguidas por las palabras con TV con 46.21 %, las palabras monosílabas con 55.41 % y, finalmente, el tipo de estímulo verbal que peor reconocimiento tuvo fueron los bisílabos con 63.43 % de errores.

### **Figura 1.**

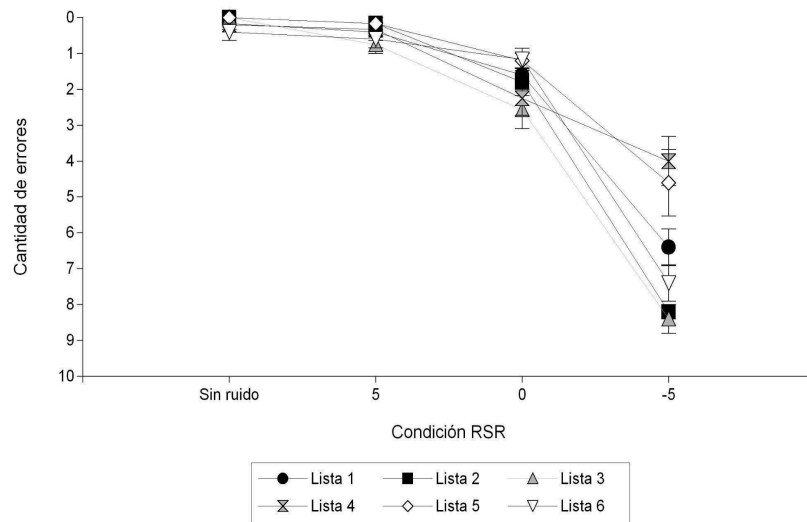
***Cantidad de errores en la discriminación de monosílabos en cada condición de RSR según listas***



El segundo estudio buscaba determinar si la complejidad en el reconocimiento de las listas era equivalente. En la Figura 1 se presenta la media de errores en la discriminación de monosílabos, para cada condición de RSR según las distintas listas. Se puede observar que, en las condiciones de RSR 5, 0 y -5, las listas 1 y 6 son las que presentan más errores de discriminación. Se realizó el test Kruskal Wallis para analizar la varianza y se observaron diferencias estadísticamente significativas entre las listas ( $p = .0186$ ); particularmente se pudo ver que las listas 2, 3, 4 y 5 no presentan diferencias significativas entre ellas, pero sí diferencias significativas respecto a las listas 1 y 6.

**Figura 2.**

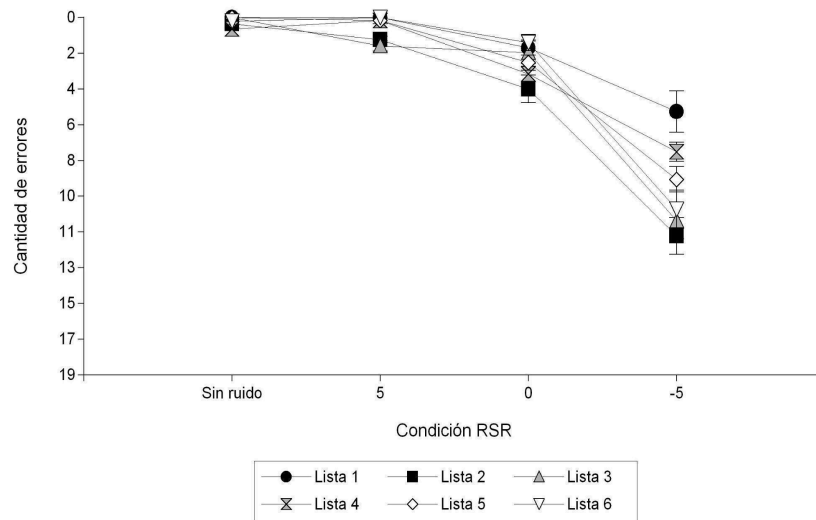
***Cantidad de errores en la discriminación de bisílabos en cada condición de RSR según listas***



En la Figura 2 se presenta la media de errores en la discriminación de bisílabos, para cada condición de RSR según las distintas listas. Se puede observar que en las condiciones sin ruido y RSR 5 hay una menor dispersión entre las distintas listas, ya que la cantidad de errores es baja. En la condición de RSR 0 las listas con más errores son la 4 y la 3; en la condición de RSR -5, las listas con más errores son las 2 y 3. Se realizó el test Kruskal-Wallis para analizar la varianza y no se observaron diferencias estadísticamente significativas entre las distintas listas ( $p = .7162$ ).

**Figura 3.**

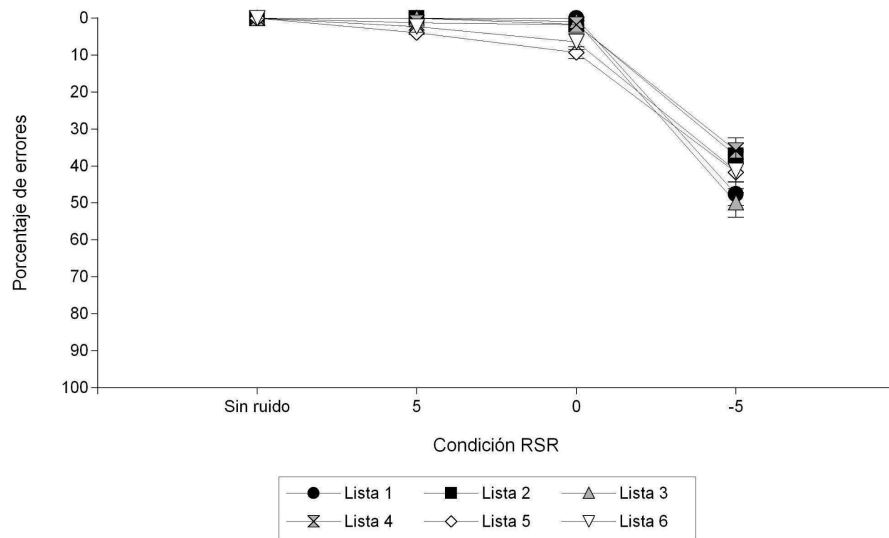
***Errores de discriminación de TV en cada condición de RSR según listas***



En la Figura 3 se presenta la media de errores en la discriminación de TV, para cada condición de RSR según las distintas listas. Se puede observar que en las condiciones sin ruido y RSR 5, hay una menor dispersión entre las distintas listas, ya que la cantidad de errores es baja. En la condición de RSR 0 las listas con más errores son la 2 y la 3; en la condición de RSR -5 las listas con más errores son las 2, 3 y 6. Se realizó el test Kruskal-Wallis para analizar la varianza y no se observaron diferencias estadísticamente significativas entre las distintas listas ( $p = .3123$ ).

**Figura 4.**

***Porcentaje de errores de la discriminación de OUGE en cada condición de RSR según listas***



En la Figura 4 se presenta la media de errores en la discriminación de OUGE, para cada condición de RSR según las distintas listas. Se puede observar que en la condición sin ruido los sujetos no presentaron errores en ninguna lista. En la condición de RSR 5 hay poca variabilidad entre las distintas listas. En la condición de RSR 0 las listas con más errores son la 5 y la 6; en la condición de RSR -5, las listas con más errores son las listas 2 y 6. Si bien en la condición de RSR -5 se observan diferencias considerables entre las listas, al realizar el test Kruskal-Wallis para analizar la varianza, al considerar todas las RSR en conjunto, no se observaron diferencias estadísticamente significativas entre las distintas listas ( $p = .1493$ ).

Para saber si los mayores niveles de error se podían asociar a un menor nivel de activación léxica de las palabras, como consecuencia de menor frecuencia de exposición en el lenguaje, se analizó la correlación entre la cantidad de errores y la frecuencia de ocurrencia en los corpus; del español NOW (Davies, 2018) y el CORPES XXI de la (Real Academia Española) y, en este último se seleccionaron solo fuentes de Argentina. En ambos corpus las palabras bisílabas y las TV obtuvieron una correlación negativa y débil, pero significativa; es decir, como se suponía, a mayor frecuencia de uso, menor error. Los valores al analizar el Corpus del Español NOW fueron  $r = -.27$  ( $p = .035$ )

para palabras bisílabas y  $r = -.26$  ( $p = .004$ ) para palabras con TV. En el caso del CORPES XXI, se obtuvo  $r = -.30$  ( $p = .018$ ) con palabras bisílabas y para palabras con TV,  $r = -.30$  ( $p = .001$ ). Para los monosílabos no se observó correlación en ninguno de los dos corpus;  $r = .08$  ( $p = .479$ ) en el Corpus NOW y  $r = .06$  ( $p = .588$ ).

Dados esos resultados, se consideró necesario analizar las diferencias en la frecuencia léxica de las palabras a nivel de cada lista. En ese cálculo se consideró, como valor de frecuencia léxica para una lista, la suma de las frecuencias léxicas de las palabras que la componían. Se aplicó el test Kruskal-Wallis para conocer si había diferencias en la varianza de frecuencia léxica de las palabras entre las distintas listas. En ningún caso se obtuvieron diferencias estadísticamente significativas. En la Tabla 4 se exponen los valores totales de frecuencia léxica de cada lista y los resultados del test.

**Tabla 4.**

*Frecuencia léxica de cada lista en el Corpus del Español NOW y el CORPES XXI, y análisis de diferencias en la frecuencia léxica entre listas para cada corpus*

Estímulo verbal	Lista	Frecuencia léxica total Corpus del Español NOW	Kruskal Wallis		Frecuencia léxica total CORPES XXI	Kruskal Wallis	
			H	p-valor		H	p-valor
Monosílabos	1	23454370	2.63	.75	65466	5.89	.31
	2	1404749			6354		
	3	8269561			41089		
	4	5981148			27502		
	5	1696216			2778		
	6	1147861			7628		
Bisílabos	1	16270222	6.59	.25	65466	9.32	.09
	2	84889359			6354		
	3	230078972			41089		
	4	45439739			27502		
	5	54667037			2778		
	6	748792475			7628		
TV	1	15468433	6.40	.26	76566	7.60	.17
	2	7074597			28661		
	3	6481829			30138		
	4	6538305			29771		
	5	2377855			28288		
	6	5061129			30624		

## Discusión

En este trabajo se analizaron los resultados del empleo de la batería BEPPA, diseñada originalmente para evaluar el reconocimiento del habla en pacientes con prótesis auditiva como instrumento para el estudio de discriminación de habla en ruido con sujetos normooyentes. En primer lugar, se grabaron las diferentes listas que componen esa batería con una locutora profesional y condiciones de estudio de alta calidad acústica. Se sintetizaron versiones para esas listas con RSR de 5, 0 y -5 dB, para el que se utilizó ruido de tipo SSN, debido a que presenta mayores niveles de enmascaramiento que el ruido blanco o rosa utilizado en la propuesta original de la batería. Se calibró la cadena de reproducción de sonidos para garantizar condiciones experimentales similares entre los diferentes sujetos y estímulos.

Con los 32 sujetos normooyentes se comparó la cantidad de errores entre las distintas condiciones de RSR evaluadas al comparar las listas y se analizó la cantidad de errores para cada palabra, al considerar todas las condiciones de RSR con el fin de identificar aquellas con un bajo nivel de reconocimiento.

En relación a las distintas condiciones de RSR se pudo observar que, de acuerdo a lo esperado, se incrementó la media de errores a medida que se disminuyó la RSR. Para los cuatro tipos de estímulos verbales (monosílabos, bisílabos, TV y OUGE) se observaron diferencias significativas ( $p < .05$ ) entre las condiciones de RSR 5 y 0, y entre RSR 0 y -5. Entre la condición sin ruido y RSR 5 solo se obtuvieron diferencias significativas para monosílabos y TV. Gurlekian et al. (2012) evaluaron la inteligibilidad en ruido de adultos con audición normal para el que se utilizaron oraciones como estímulo verbal y encontraron resultados similares a los presentados en esta investigación, ya que para RSR de 20, 15 10 y 5 dB no observaron diferencias significativas, pero que sí las había respecto a RSR 0 y -5 dB. Cabe destacar que para

RSR -5 dB encontraron un 40 % de aciertos en discriminación y, en esta investigación, para OUGE se ha reportado un 70 % de aciertos. Esta notoria diferencia se puede atribuir al ambiente acústico en el que se realizaron las evaluaciones. En la investigación de Gurlekian et al. (2012), la evaluación de discriminación se realizó en espacios con ruido ambiente de 40-50 dB, mientras que en este estudio se lo efectuó dentro de una cabina sonoamortiguada. Otra diferencia que puede explicar las diferencias obtenidas entre ambos trabajos está dada por la forma de contabilizar los errores. Mientras que en el trabajo de Gurlekian et al. (2012) se considera errónea la totalidad de la oración si la persona repite mal alguna de las palabras de contenido, en el caso de esta investigación se contabilizó el porcentaje de la oración que se repitió erróneamente, lo que contribuye a disminuir la proporción de errores.

Wilson y Cates (2008) aplicaron el test WIN, compuesto por listas de monosílabos y ruido conversacional multihablante, y fijaron la señal de ruido en 60 dB y variaron el nivel de habla para obtener RSR entre 24 y 0 dB. Las señales fueron entregadas de forma binaural en listas de 10 palabras para cada condición de RSR; los participantes tenían audición normal y edades comprendidas entre 19 y 29 años. Su condición de RSR 0 dB sería equivalente a la del presente trabajo, tanto la señal de habla como el ruido a 60 dB. Sin embargo, en sus resultados encontraron que en RSR 0 solo un 33.8 % de reconocimiento correcto de monosílabos, mientras que, en este trabajo, los resultados muestran valores muy superiores con 78 % de aciertos en RSR 0 dB. En otro estudio Wilson et al. (2008) evaluaron la DHR al utilizar monosílabos de tres sets de listas distintas y SSN fijo a 72 dB. Si bien no utilizaron los mismos niveles de RSR, al saber que en RSR -2 dB obtuvieron valores de aciertos de 28 % - 36 % y en 3 dB de RSR de 63 % - 72 %, se pudo inferir que su resultado en una RSR de 0 dB se encontraría dentro de ese margen, por lo que los resultados de esta investigación continúan siendo

mayores. La gran diferencia aquí está dada por las distintas lenguas de los estudios. Sería ideal poder comparar con resultados en lengua española, pero no se han podido encontrar antecedentes en español que analicen la discriminación de monosílabos en ruido en detalle y condiciones similares. En inglés se utilizan típicamente palabras monosilábicas en las pruebas de discriminación del habla, mientras que en español es más común el uso de palabras bisilábicas para evaluar esta función (Shi, 2014). Esto se podría atribuir a que en español los monosílabos son más escasos y, en su mayoría, son palabras de función (artículos, pronombres, preposiciones), mientras que en inglés hay gran frecuencia de palabras monosilábicas, muchas de las cuales son palabras de contenido (García Lecumberri, 2016).

En la presente investigación, al comparar los resultados entre los distintos tipos de estímulos verbales, se pudo observar que en general a menor contexto del estímulo hubo un mayor deterioro en la tasa de reconocimiento al aumentar el ruido de fondo. Es decir que, en general, la OUGE tuvo un mejor desempeño, seguidas por las palabras con TV, bisílabos y, por último, monosílabos, a excepción de la condición RSR -5, en la que los bisílabos tuvieron un peor desempeño que los monosílabos. Es esperable que las oraciones sean las que tengan mejor reconocimiento, ya que presentan mayor información contextual, por lo tanto, si un sujeto pierde parte del mensaje, puede fácilmente recuperarlo a través del conocimiento lingüístico (Fuente et al., 2011). En relación al peor desempeño de bisílabos en RSR -5, es necesario considerar que las listas de palabras bisílabas están compuestas por 10 estímulos frente a las listas de monosílabos que están compuestas por 15 estímulos que puede llevar a que cada error tenga un impacto mayor sobre el porcentaje total de errores.

Respecto a la comparación entre listas, los resultados mostraron que las mayores diferencias se encuentran en la condición de RSR -5 y, probablemente, se deba a que al

ser la condición más difícil es la que permite ver mayores variaciones. En el caso de los bisílabos, las listas con más errores para la condición de RSR -5 fueron la 2 y 3. Se analizaron las diferencias entre las listas de bisílabos, para las que se consideraron todas las condiciones de RSR, y no se observaron diferencias significativas ( $p > .05$ ). Por su parte, las listas de monosílabos 1 y 6 son las que presentaron más errores de discriminación, tanto en RSR 0 como en RSR -5, y se constató que las listas de monosílabos 2, 3, 4 y 5 no presentaban diferencias significativas entre ellas, mientras que sí las presentaban respecto a las listas 1 y 6. En relación a las TV, las listas tuvieron una cantidad de errores similares y no se observaron diferencias significativas entre las distintas listas. Finalmente, respecto a las listas de OUGE, se observó que en RSR -5 la lista 2 tuvo un desempeño menor que el resto de las listas, seguida por la lista 6; sin embargo, al analizar las diferencias entre listas, al considerar todas las condiciones, no se observaron diferencias significativas. Para evaluar a una persona con la metodología aquí propuesta son necesarias al menos cuatro listas distintas y en la mayoría de los estímulos verbales no se observaron diferencias significativas entre las listas, lo que indicaría que se podrían seleccionar cuatro de las seis listas disponibles sin riesgo de caer en errores sistemáticos. La excepción a esto serían los monosílabos; con este estímulo verbal sería conveniente utilizar específicamente las listas 2, 3, 4 y 5. Cabe destacar que esto aplica, particularmente, para los audios generados por el equipo que participó en esta investigación y la cadena de medición implementada.

Al analizar la correlación entre la cantidad de errores y la frecuencia léxica de las palabras, se observó una correlación negativa débil significativa en el caso de las palabras bisílabas y las palabras con TV, tanto en el Corpus del Español NOW (Davies, 2018) como en el CORPES XXI. Es decir que aquellas palabras con menos frecuencia de ocurrencia tuvieron cierta tendencia a presentar más errores. Según esto, es

importante que las distintas listas no tengan diferencias en la frecuencia léxica de sus palabras, ya que una lista compuesta por mayor cantidad de palabras con gran frecuencia léxica correría con cierta ventaja de identificación. Al comparar las distintas listas, con cierta frecuencia léxica, no se observaron diferencias significativas, por lo que se considera que las listas son equivalentes en este aspecto para evaluar la discriminación del habla en ruido.

En relación a las limitaciones de este trabajo, se puede mencionar que el tamaño de la muestra es acotado. A su vez, por los resultados obtenidos, sería interesante que en futuros trabajos de DHR en personas con audición normal se pudieran incluir algunas pistas con RSR menores.

### **Conclusión**

En este trabajo se realizó un análisis inicial de un material preexistente en español rioplatense que permite evaluar la discriminación del habla en ruido. Al generar una versión propia en la que se utiliza ruido más competitivo que el originalmente propuesto, se pudo observar que para palabras bisílabas, TV y OUGE, las listas analizadas resultaron equivalentes respecto a la cantidad de errores para cada RSR. Distinto fue el caso de las palabras monosílabas en las que hubo dos de las seis listas que obtuvieron menores valores con diferencias estadísticamente significativas. En relación a la frecuencia léxica, las listas de palabras bisílabas, monosílabas y TV, también resultaron equivalentes, ya que no se detectaron diferencias estadísticamente significativas. De acuerdo a los aspectos estudiados, se concluye que esta batería resulta de utilidad para evaluar sujetos normooyentes en la clínica audiológica.

La discriminación del habla en ruido cobra cada vez más importancia en el área de la audiológica, ya que cada vez son más las personas que manifiestan dificultades en la discriminación del habla en ruido y que presentan audición dentro de parámetros

normales. Esta dificultad podría ser un indicador temprano de deterioro auditivo y su evaluación se acerca más a los contextos de escucha diaria de las personas. Por tal motivo, resulta fundamental conocer las características de los materiales disponibles en la región para su evaluación, como así también estandarizar los valores que se obtienen en distintas poblaciones para que, posteriormente, pueda ser incluido en la batería audiológica de rutina, no solo para personas con prótesis auditivas, sino también para personas con audición normal o hipoacusia leve.

### Referencias

- Armstrong, N. M., Oosterloo, B. C., Croll, P. H., Ikram, M. A. y Goedegebure, A. (2020). Discrimination of degrees of auditory performance from the digits-in-noise test based on hearing status. *International Journal of Audiology*, 59(12), 897-904. <https://doi.org/10.1080/14992027.2020.1787531>
- Aronson, L., Milone, D., Martínez, C., Estienne, P., Tomassi D., Rufiner H. L. y Torres, M. E. (2007). Batería para la evaluación del reconocimiento del habla en pacientes con prótesis auditiva. *Revista de la Federación Argentina de Sociedades de Otorrinolaringología*, 14(1), 17-24. <https://sinc.unl.edu.ar/sinc-publications/2007/AEMMTRT07/>
- Bargetto Fernández, M. Á. y Riffo Ocares, B. (2019). El reconocimiento de palabras y el acceso léxico: revisión de modelos y pruebas experimentales. *Boletín de Filología*, 54(1), 341-361. <https://doi.org/10.4067/S0718-93032019000100341>
- Beck, D. L., Danhauer, J. L., Abrams, H. B., Atcherson, S. R., Brown, D. K., Chasin, M., Greer Clark, J., De Placido, C., Edwards, B., Fabry, D., Flexer, C., Fligor, B., Frazer, G., Galster, J., Gifford, L., Johnson, C. E., Madell, J., Moore, D. R., Roeser, R. J., Saunders, G. H., Searchfield, G. D., Spankovich, C., Valente, M. y Wolfe, J. (2018). Audiologic considerations for people with normal hearing

sensitivity yet hearing difficulty and/or speech-in-noise problems. *Hearing Review*, 25(10), 28-38.

<https://hearingreview.com/hearing-loss/patient-care/evaluation/audiologic-considerations-people-normal-hearing-sensitivity-yet-hearing-difficulty-andor-speech-noise-problems>

Brookes, M. (2009) VoiceBOX: speech processing toolbox for MATLAB.

<http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>

Carlo, M. A., Wilson, R. H. y Villanueva-Reyes, A. (2020). Psychometric

Characteristics of Spanish Monosyllabic, Bisyllabic, and Trisyllabic Words for Use in Word-Recognition Protocols. *Journal of the American Academy of Audiology*, 31(07), 531-546. <https://doi.org/10.1055/s-0040-1709446>

Cox, R. M., Alexander, G. C. y Gilmore, C. (1987). Development of the connected speech test (CST). *Ear and Hearing*, 8(5), 119S-126S.

<https://doi.org/10.1097/00003446-198710001-00010>

Davies, M. (2018). *News on the Web Corpus (NOW)*. Brigham Young University.

<https://www.corpusdelespanol.org/now/>

Di Rienzo, J. A., Casanoves, F., Balzarini, M. G., Gonzalez, L., Tablada, M. y Robledo, C. W. (2020). InfoStat versión 2020. <http://www.infostat.com.ar>

Fuente, A., Hormazábal, X., López, A. y Bowen, M. (2011). Efecto de supresión eferente de las emisiones otoacústicas transientes y discriminación de habla en ruido. *Revista Chilena de Fonoaudiología*, 10, 7-16.

<https://doi.org/10.5354/rcdf.v10i0.17349>

Gaetán, S., Tenutta, M., Bertinatti, A., Vicente, F., Muratore, J., Maggi, A. L., Hinalaf, M. y Ferreyra, S. P. (2019). Evaluación de Características Acústicas de Cabinas

- Audiométricas Fijas y Móviles del CINTRA. *Mecánica Computacional*, 37(5), 67-76. <https://amcaonline.org.ar/ojs/index.php/mc/article/view/5776>
- García Lecumberri, M. L., Barker, J., Marxer, R. y Cooke, M. (2016). *Language Effects in Noise-Induced Word Misperceptions*. INTERSPEECH. September 8–12, San Francisco, USA. <https://doi.org/10.21437/Interspeech.2016-330>
- Gurlekian, J. A., Babnik, E. y Torres, H. M. (2008). Desarrollo de una prueba de inteligibilidad de habla en ambientes ruidosos para niños en edad escolar. *Revista de Logopedia, Foniatría y Audiología*, 28(3), 138-148. [https://doi.org/10.1016/S0214-4603\(08\)70052-4](https://doi.org/10.1016/S0214-4603(08)70052-4)
- Gurlekian, J., Pavoni, A., López, A. y Lascano, M. J. (2012, enero). Trabajo presentado en 32º Congreso de ORL y actividades conexas. Rosario. Argentina. [https://www.researchgate.net/publication/289996700\\_Evaluacion\\_rapida\\_de\\_la\\_inteligibilidad\\_de\\_habla\\_en\\_ruido](https://www.researchgate.net/publication/289996700_Evaluacion_rapida_de_la_inteligibilidad_de_habla_en_ruido)
- Holder, J. T., Levin, L. M. y Gifford, R. H. (2018). Speech recognition in noise for adults with normal hearing: Age-normative performance for AzBio, BKB-SIN, and QuickSIN. *Otology & Neurotology*, 39(10), e972. <https://doi.org/10.1097/MAO.0000000000002003>
- IRAM, 4026:1986. Cabinas audiométricas. Instituto Argentino de Normalización y Certificación.
- Kalikow, D. N., Stevens, K. N. y Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, 61(5), 1337-1351. <https://doi.org/10.1121/1.381436>
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J. y Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio

loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 116(4), 2395-2405.

<https://doi.org/10.1121/1.1784440>

Killion, M. C. y Villchur, E. (1993). Kessler was right-partly: But SIN test shows some aids improve hearing in noise. *Hearing Journal*, 46, 31-31.

<https://www.etymotic.com/research-publications/>

Koole, A., Nagtegaal, A. P., Homans, N. C., Hofman, A., Baatenburg de Jong, R. J. y Goedegebure, A. (2016). Using the digits-in-noise test to estimate age-related hearing loss. *Ear and Hearing*, 37(5), 508-513.

<https://doi.org/10.1097/AUD.0000000000000282>

Luce, P. A. y Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19(1), 1-36.

<https://doi.org/10.1097/00003446-199802000-00001>

Maggio, M., Marrero-Aguiar, V. M. y Calvo, J. C. (2016). Material para la evaluación de la percepción del habla en ruido en niños: frases PIP-UNED. *FIAPAS: Federación Ibérica de Asociaciones de Padres y Amigos de los Sordos*, (158), 22-23.

<https://www.phonak-pip.es/download/material-para-la-evaluacion-de-la-percepcion-del-habla-en-ruido-en-ninos-frases-pip-uned/>

Marrero-Aguiar, V. (2015). La percepción del habla en ruido. Estudio experimental sobre una aplicación para la evaluación audiológica infantil. *Revista Española de Lingüística*, 45(1), 129-151.

<http://revista.sel.edu.es/index.php/revista/article/view/94>

Nilsson, M., Soli, S. D. y Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise.

*The Journal of the Acoustical Society of America*, 95(2), 1085-1099.

<https://doi.org/10.1121/1.408469>

Real Academia Española. Banco de datos (CORPES XXI) [en línea]. Corpus del Español del Siglo XXI (CORPES).

<https://apps2.rae.es/CORPES/view/inicioExterno.view>

Rönnberg, J., Rudner, M., Lunner, T. y Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health*, 12(49), 263-269. <https://doi.org/10.4103/1463-1741.70505>

Shi, L. F. (2014). Validating models of clinical word recognition tests for Spanish/English bilinguals. *Journal of Speech, Language, and Hearing Research*, 57(5), 1896-1907. [https://doi.org/10.1044/2014\\_JSLHR-H-13-0138](https://doi.org/10.1044/2014_JSLHR-H-13-0138)

Soli, S. D., Vermiglio, A., Wen, K. y Filesari, C. A. (2002). Development of the hearing in noise test (HINT) in Spanish. *The Journal of the Acoustical Society of America*, 112(5), 2384-2384. <https://doi.org/10.1121/1.4779699>

Taylor, B. (2003). Speech-in-noise tests: How and why to include them in your basic test battery. *The Hearing Journal*, 56(1), 40-42. <https://doi.org/10.1097/01.HJ.0000293000.76300.ff>

Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology*, 14(9), 453-470. <https://doi.org/10.1055/s-0040-1715938>

Wilson, R. H., Carnell, C. S. y Cleghorn, A. L. (2007). The Words-in-Noise (WIN) Test with Multitalker Babble and Speech-Spectrum Noise Maskers. *Journal of the American Academy of Audiology*, 18(06), 522-529. <https://doi.org/10.3766/jaaa.18.6.7>

- Wilson, R. H. y Cates, W. B. (2008). A comparison of two word-recognition tasks in multitalker babble: Speech Recognition in Noise Test (SPRINT) and Words-in-Noise Test (WIN). *Journal of the American Academy of Audiology*, 19(07), 548-556. <https://doi.org/10.3766/jaaa.19.7.4>
- Wilson, R. H. y McArdle, R. A. (2008). Speech-in-noise measures as necessary components of routine audiologic evaluation and auditory processing disorder evaluation. En A. T. Cacace y D. J. McFarland (Eds.), *Current controversies in central auditory processing disorder* (pp. 151-168). Plural Publishing.
- Wilson, R. H., McArdle, R. y Roberts, H. (2008). A Comparison of Recognition Performances in Speech-Spectrum Noise by Listeners with Normal Hearing on PB-50, CID W-22, NU-6, W-1 Spondaic Words, and Monosyllabic Digits Spoken by the Same Speaker. *Journal of the American Academy of Audiology*, 19(06), 496-506. <https://doi.org/10.3766/jaaa.19.6.5>

Recibido: 21 de junio de 2022

Aceptado: 24 de abril de 2023