

Otología y Otoneurología

Desarrollo de listas psicométricamente equivalentes en español rioplatense para pruebas de Porcentaje de Reconocimiento de Palabras

Homenaje al Prof. Dr. Juan Manuel Tato

Development of psychometrically equivalent lists in rioplatense spanish for Word Recognition Score Test

Tribute to Prof. Dr. Juan Manuel Tato

Desenvolvimento de listas psicometricamente equivalentes em espanhol rioplatense para testes de Indice Percentual de Reconhecimento de Fala

Homenagem ao Prof. Dr. Juan Manuel Tato

Ing. Andrés Piegari⁽¹⁾, Ing. Horacio Cristiani⁽²⁾, Lic. Sabrina Alonso⁽²⁾, Lic. Ornella Virgallito⁽²⁾, Ing. Matías Pace⁽¹⁾, Ing. Patricio Bertato⁽¹⁾, Dra. Verónica Del Vecchio⁽³⁾, PhD. Sebastián Ausili⁽⁴⁾

Resumen

Introducción: Se desarrolló un conjunto de listas de palabras psicométricamente equivalentes y con balance fonético para ser utilizadas en pruebas de porcentaje de reconocimiento de palabras en la variante del español rioplatense. Este material se compone de 10 listas de 25 palabras cada una. Fue grabado digitalmente por un hablante de sexo femenino y validado en términos de equivalencia psicométrica.

Material y Método: Un primer conjunto de 300 palabras fue presentado a 24 sujetos, divididos en dos grupos con audición normal (edad promedio: 30 años) con niveles de presentación en el rango de -5 a 40 dB HL en pasos de 5 dB. Luego de la recolección

de datos, las 250 palabras con mejor reconocimiento se distribuyeron en 10 listas. En esta distribución, se priorizó la selección de palabras de manera que se lograra la equivalencia psicométrica entre las listas conformadas y un buen grado de balance fonético.

Resultados: Las curvas psicométricas correspondientes a las listas fueron obtenidas con base en los resultados registrados. Se verificó, visualmente y mediante pruebas no paramétricas, que no existen diferencias estadísticamente significativas entre las listas. Para el análisis del balance fonético de las listas, se emplearon distintas herramientas estadísticas que mostraron un alto grado de representación fonética con respecto al corpus principal de palabras.

⁽¹⁾ Universidad Nacional de Tres de Febrero, Buenos Aires, Argentina.

⁽²⁾ Mutualidad Argentina de Hipoacúsicos, Buenos Aires, Argentina.

⁽³⁾ Asociación Argentina de Audiología, Buenos Aires, Argentina.

⁽⁴⁾ University of Miami, Department of Otolaryngology, Ear Institute, Miami, USA.

Mail de contacto: apiegari@untref.edu.ar

Fecha de envío: 21 de septiembre de 2024 - Fecha de aceptación: 26 de septiembre de 2024.

Conclusión: Este material de habla, desarrollado específicamente para la obtención de porcentajes en el reconocimiento de palabras, ha sido validado en cuanto a uniformidad psicométrica y balance fonético de las listas que lo conforman y se encuentra disponible para los profesionales que lo requieran.

Palabras clave: logaudiometría, balance fonético, pruebas de habla, español rioplatense, porcentaje de reconocimiento de palabras.

Abstract

Introduction: A set of psychometrically equivalent and phonetically balanced word lists was developed for use in Word Recognition Score tests in the Rioplatense Spanish variant. This material consists of 10 lists, each containing 25 words. The lists were digitally recorded by a female speaker and validated in terms of psychometric equivalence.

Material and Method: An initial set of 300 words was presented to 24 subjects, divided into two groups with normal hearing (average age: 30 years). Presentation levels ranged from -5 to 40 dB HL in 5 dB steps. After data collection, the 250 words with the highest recognition rates were distributed into 10 lists. During this distribution, priority was given to word selection to ensure psychometric equivalence between the lists and a high degree of phonetic balance.

Results: Psychometric curves corresponding to the lists were obtained based on the recorded results. It was verified, visually and through non-parametric tests, that there were no statistically significant differences between the lists. For the phonetic balance analysis of the lists, various statistical tools were employed, which demonstrated a high degree of phonetic representation in relation to the primary word corpus.

Conclusion: This speech material, specifically developed for obtaining word recognition scores, has been validated in terms of psychometric uniformity and phonetic balance of the lists, and is available for professionals who may require it.

Keywords: speech audiometry, phonetic balance, speech tests, Rioplatense Spanish, word recognition score.

Resumo

Introdução: Um conjunto de listas de palavras psicometricamente equivalentes e com equilíbrio fonético foi desenvolvido para ser utilizado em testes de escores de reconhecimento de palavras na variante do espanhol rioplatense. Este material é composto por 10 listas, cada uma contendo 25 palavras. As

listas foram gravadas digitalmente por uma falante do sexo feminino e validadas em termos de equivalência psicométrica.

Material e Método: Um conjunto inicial de 300 palavras foi apresentado a 24 sujeitos, divididos em dois grupos com audição normal (idade média: 30 anos), com níveis de apresentação variando de -5 a 40 dB HL em incrementos de 5 dB. Após a coleta de dados, as 250 palavras com melhor reconhecimento foram distribuídas em 10 listas. Durante esta distribuição, priorizou-se a seleção de palavras para alcançar equivalência psicométrica entre as listas e um bom grau de equilíbrio fonético.

Resultados: As curvas psicométricas correspondentes às listas foram obtidas com base nos resultados registrados. Verificou-se, visualmente e por meio de testes não paramétricos, que não existem diferenças estatisticamente significativas entre as listas. Para a análise do equilíbrio fonético das listas, foram utilizadas várias ferramentas estatísticas, que demonstraram um alto grau de representação fonética em relação ao corpus principal de palavras.

Conclusão: Este material de fala, desenvolvido especificamente para a obtenção de escores de reconhecimento de palavras, foi validado em termos de uniformidade psicométrica e equilíbrio fonético das listas que o compõem, e está disponível para os profissionais que o necessitarem.

Palavras-chave: logaudiometria, equilíbrio fonético, testes de fala, espanhol rioplatense, escore de reconhecimento de palavras.

Introducción

El lenguaje desempeña un papel fundamental, no sólo como medio de comunicación, sino también como un elemento que contribuye al desarrollo cognitivo del ser humano. Es una herramienta para la educación y la base de las relaciones sociales. La hipoacusia es un trastorno sensorial que se manifiesta como una reducción en la capacidad auditiva. Es una condición común que afecta a personas en todo el mundo, independientemente de su ubicación geográfica.

La pérdida auditiva, según su grado y tipo, puede impactar significativamente sobre los procesos de comunicación y, en consecuencia, marginar al individuo de esta habilidad esencial. Una evaluación audiológica completa es esencial para la adecuada intervención terapéutica de las personas con discapacidad auditiva. En este sentido, las pruebas de percepción y reconocimiento del habla se encuentran entre las más importantes para el diagnóstico audiológico⁽¹⁾. Combinada con la audiometría tonal,

constituyen las herramientas más empleadas para determinar el tipo y grado de pérdida auditiva.

Las pruebas de reconocimiento del habla tienen como objetivo medir la capacidad de una persona para discriminar el lenguaje hablado, lo que puede facilitar el diagnóstico topográfico de posibles lesiones en la vía auditiva. Estas pruebas desempeñan un papel crucial al evaluar la capacidad de comunicación del sujeto en situaciones de la vida cotidiana. Además, son fundamentales en la adaptación de audífonos o implantes cocleares como una medida estándar para evaluar el beneficio del dispositivo de ayuda auditiva. También son útiles para evaluar la posibilidad de mejoras postquirúrgicas, además de detectar posibles casos de simulación de hipoacusia.

A pesar de la diversidad de tipos pruebas de reconocimiento de palabras, todas comparten el objetivo de evaluar la capacidad de comprensión del sujeto mediante su respuesta verbal ante la presentación de un estímulo a niveles específicos de intensidad. Dentro de este tipo de pruebas, se destacan tres medidas fundamentales: el umbral de detección del habla (SDT, Speech Detection Threshold, por sus siglas en inglés), el umbral de reconocimiento del habla (SRT, Speech Recognition Threshold, por sus siglas en inglés) y el porcentaje de reconocimiento de palabras (WRS, Word Recognition Score, por sus siglas en inglés). Tanto el SDT como el SRT buscan determinar un valor de umbral, es decir, el mínimo nivel de intensidad necesario para obtener un tipo específico de respuesta. El SDT se enfoca en encontrar el umbral en el cual un sujeto detecta la presencia del estímulo de habla, pero no logra comprender la palabra. En el caso del SRT, la búsqueda del umbral se centra en encontrar el nivel de intensidad mínimo para el cual el sujeto es capaz de repetir correctamente el 50% de las palabras presentadas. Por último, el WRS indica el porcentaje de reconocimiento obtenido para determinados niveles de presentación, generalmente superiores al SRT, en otras palabras, supraliminales. El material utilizado para obtener el WRS es, en muchas lenguas del mundo, diferente al empleado para hallar el SRT. En un trabajo anterior, se describió la generación y validación de un conjunto de palabras psicométricamente equivalentes aptas para la obtención del SRT en el español rioplatense⁽²⁾. El presente trabajo muestra el desarrollo y validación de material destinado a la obtención del WRS.

Para obtener validez y confiabilidad en las pruebas de WRS, es fundamental considerar los siguientes criterios durante el desarrollo de nuevo mate-

rial: 1) elegir palabras de uso corriente, 2) equilibrar fonéticamente cada lista y 3) mantener un rango de dificultad uniforme entre listas. También, debe considerarse el uso de palabras con menor cantidad de sílabas que las empleadas para SRT, esto es, material que presente mayor dificultad de reconocimiento y con curvas psicométricas con menor pendiente⁽³⁾. Además, una vez seleccionado el material de prueba, la grabación de las listas debe ser realizada por hablantes nativos que exhiban un acento estándar del idioma regional. Esto se debe a que las variaciones dialécticas entre los hablantes de un mismo idioma pueden condicionar los resultados de reconocimiento, especialmente en niveles bajos de presentación⁽⁴⁾.

Diferentes materiales para pruebas de palabras se han desarrollado en inglés. Son conocidas las listas de palabras monosilábicas CID Auditory Test W-22⁽⁵⁾ y Northwestern University No.6⁽⁶⁾ utilizadas para la obtención del WRS y las listas de palabras espondeicas (palabras compuestas) para obtener SRT, denominadas CID W-1. Puede encontrarse material grabado y validado en muchos otros idiomas tales como el ruso⁽⁴⁾, el alemán⁽⁷⁾, el chino mandarín⁽⁸⁾ y el griego⁽⁹⁾, entre otros. Desde su presentación, a fines de la década de 1940, el material utilizado mayoritariamente para la evaluación de reconocimiento de palabras en Argentina es el desarrollado por el Dr. Tato et al.⁽¹⁰⁾. El producto de este trabajo, pionero a nivel mundial, es un conjunto de palabras fonéticamente balanceadas que fueron utilizadas en Argentina y otros países de habla hispana durante casi 80 años. Este material se compone de 12 listas de 25 palabras bisilábicas y acentuación grave. Es importante destacar la calidad de este material y el procedimiento minucioso para generarlo, a pesar de los escasos medios tecnológicos de la época. Sin embargo, al ser listas con varias décadas desde su creación, incluyen ciertas palabras poco familiares para los sujetos evaluados en la actualidad; algunas por obsolescencia y otras por simple desconocimiento. Esto conlleva el problema de no poder determinar si los errores en la repetición de esas palabras tienen lugar a causa de un problema auditivo o por desconocimiento de la palabra en cuestión. Palabras como «ledo», «jade», «jaspe» y «luso» difícilmente sean repetidas en condiciones donde otras como «moldes», «menta» o «Carlos» sean fácilmente reconocibles y repetidas. Un trabajo del año 2012⁽¹¹⁾ advirtió sobre las diferencias y las posibilidades de cometer errores en la repetición de ciertas palabras de este material. Por otro lado, las funciones psicométricas de cada una de las listas no han sido publicadas. Esta ausencia de validación impide co-

robocar si las listas poseen un grado de dificultad de reconocimiento equivalente. La carencia de investigaciones, sumado al consenso existente entre los profesionales de nuestro medio sobre los resultados poco confiables de las pruebas realizadas con este material, motivaron la realización de este trabajo. Por lo tanto, surge la necesidad de desarrollar un nuevo conjunto de listas de palabras para pruebas de WRS en español rioplatense, inspirado en el desarrollo del Prof. Tato, pero con las ventajas tecnológicas del manejo de grandes cantidades de datos. El presente trabajo se basó en los criterios de desarrollo mencionados para la creación de un nuevo material de prueba. En una primera fase, se estableció la frecuencia de aparición de los fonemas en el idioma español rioplatense, utilizando un corpus generado a través de la recopilación de diferentes fragmentos de comunicación verbal. Posteriormente, se determinó la familiaridad de las palabras para su uso en la comunicación con base en el juicio de hablantes nativos. Luego de seleccionar y grabar las palabras, se llevó a cabo el proceso de validación de estas mediante un grupo de sujetos con audición normal para finalmente confeccionar las listas de palabras verificando sus características psicométricas y balance fonético.

Material y Método

Participantes

Un grupo de 24 hablantes nativos del español rioplatense participaron en este estudio (13 mujeres) con una edad promedio de 30 años ($SD = 6$ años). El rango de edades abarcó desde los 19 a los 41 años. Todos los participantes presentaron audición normal, con umbrales de audición por conducción aérea menores a 15 dB HL en el rango de frecuencias estándar de 125 a 8000 Hz. Para la evaluación del material, se seleccionó el oído con mejor promedio de tonos puros para las frecuencias 500, 1000 y 2000 Hz (PTA, Pure Tone Average, por sus siglas en inglés). Todos los procedimientos experimentales fueron aprobados por el comité de ética de la Universidad Nacional de Tres de Febrero (UNTREF 2023-0421-4).

Material de habla

El material de habla que sirvió como base para la selección de palabras a utilizar en las listas consistió en diversas fuentes extraídas de videos disponibles en una plataforma de contenido en línea (YouTube, Google LLC). Se seleccionaron videos provenientes de diversos canales que cumplieran con criterios de calidad en la relación señal/ruido y diversidad en

el origen y género, evitando aquellos con contenido musical o fuerte sesgo regional⁽¹²⁾. Se eligieron diferentes canales de YouTube para obtener una mezcla de lenguajes formales e informales, programas para niños y adultos. La mayor parte de las palabras provinieron de videos infantiles, asumiendo que su lenguaje es conocido por todos los adultos. El conjunto inicial consistió en 23.017.107 palabras. A partir de este corpus, se llevaron a cabo dos tareas: por un lado, se realizó el análisis fonético para determinar la frecuencia de aparición de cada fonema en el español rioplatense. La segunda tarea fue la separación de las palabras bisílabas con acentuación grave. Para el análisis fonético, se empleó una herramienta de código abierto llamada Epitran⁽¹³⁾, la cual permite traducir palabras escritas a fonemas mediante la configuración de reglas fonéticas específicas. Durante este proceso, surgieron algunos errores debido a limitaciones del *software* o de las transcripciones de YouTube. Estos errores, que afectaron al 0.17% del total analizado, fueron descartados y con el remanente se logró realizar la traducción identificando 99.716.646 fonemas. Posteriormente, se calculó la frecuencia de aparición de cada fonema para el español rioplatense. Los resultados obtenidos son similares a los publicados por diversos autores⁽¹⁴⁾. Posteriormente, se seleccionaron las palabras bisílabas con acentuación grave utilizando un *software* de código abierto llamado Pylabeador⁽¹⁵⁾, el cual fue modificado para ajustarse a las reglas de separación silábica del español. Este *software* permitió clasificar las palabras según el número de sílabas. Sólo se conservaron las palabras bisílabas de acentuación grave, resultando en 5.636.925 palabras. De este conjunto se separaron las palabras que presentaron más de 1200 repeticiones, que resultaron ser 561, luego de eliminar nombres propios y homófonos. Estas palabras fueron evaluadas por un conjunto independiente de fonoaudiólogos en términos de actualidad y familiaridad. Se seleccionaron finalmente 300 palabras que representaban mejor la distribución fonética del español rioplatense. Las 300 palabras fueron grabadas para las etapas posteriores del proyecto.

Grabación y procesamiento del material

Las grabaciones se realizaron en una cabina de sonido de doble pared, tratada acústicamente en la Universidad Nacional de Tres de Febrero. La locutora fue una profesional hablante nativa del español rioplatense, lo que aseguró la autenticidad y la calidad lingüística del material. Se utilizó un micrófono

M50 (marca Earthworks Inc.), ubicado aproximadamente a 15 cm frente a la locutora, con un filtro antipop entre ambos. La señal fue amplificada y digitalizada usando un convertidor analógico-digital Fireface UCX (marca RME). Las grabaciones se realizaron con una frecuencia de muestreo de 44.1 kHz y una cuantificación de 24 bits. Durante todas las sesiones, se solicitó a la locutora que emitiera cada palabra tres veces. Un juez nativo calificó las repeticiones de cada palabra según la calidad percibida de la producción, y se seleccionó la más sobresaliente para incluirla en las etapas siguientes del estudio. Cualquier palabra que se considerara mal grabada debido a la distorsión por picos, ruido o cualquier otra anomalía se volvió a grabar o se eliminó del conjunto. Se generó una señal de calibración de 1 kHz y su intensidad se tomó como referencia para obtener un nivel de 0 dB VU en un vúmetro. Después del proceso de calificación, el valor eficaz (o RMS, Root Mean Square, por sus siglas en inglés) de cada palabra se ajustó para que la deflexión promedio coincidiera con la correspondiente al tono de calibración en el vúmetro, según lo indicado por la norma ANSI S3.6-2010⁽¹⁶⁾.

Procedimiento

Se utilizó un *software* de producción propia, basado en MATLAB (versión R2022a, The MathWorks Inc.), para controlar la presentación de los estímulos y almacenar los resultados. La señal de audio se envió desde la computadora a la entrada externa del audiómetro AC40 (marca Interacoustics) y se presentó a través de auriculares TDH-39 (marca Telephonics). La prueba se llevó a cabo en una cabina sonoamortiguada de doble pared, la cual cumple con la norma ANSI/ASA S3.1-1999 (R2018)⁽¹⁷⁾ para niveles de ruido ambiente máximos permitidos. Todas las palabras se presentaron en el oído con el PTA más bajo. Antes de probar a cada sujeto, se ajustó la sensibilidad de entrada del audiómetro a 0 dB VU utilizando el tono de calibración de 1 kHz. Además, la calibración se verificó semanalmente durante el período de adquisición de datos, de acuerdo con la norma ANSI S3.6-2010. Los participantes no estaban familiarizados con las palabras antes de la prueba. Los niveles de presentación variaron entre -5 y 40 dB HL en pasos de 5 dB. De las 300 palabras bisilábicas del corpus seleccionado, se tomaron 150 aleatoriamente y se distribuyeron en 6 listas de 25 palabras cada una. Estas listas fueron evaluadas por la mitad de los oyentes en los 10 niveles de presentación mencionados. A continuación, se tomaron las 150 palabras restantes y se reagruparon aleo-

riamente en 6 listas, de 25 palabras cada una, para ser evaluadas por la otra mitad de los sujetos. A los sujetos simplemente se les instruyó repetir la palabra que escuchaban.

Análisis de datos

Todas las palabras preseleccionadas fueron presentadas a 12 participantes en cada uno de los 10 niveles. Luego se obtuvo la cantidad total de aciertos para todos los niveles de presentación (total global) junto con el porcentaje de aciertos para cada nivel. Basados en el total global, se asignó un puntaje a cada palabra. Las palabras se ordenan en función de este puntaje, de mayor a menor. Luego, se descartaron las 50 palabras con menor puntaje para formar 10 conjuntos de 25 palabras cada uno.

Para el proceso de generación de las listas definitivas, la primera palabra de la lista ordenada se asignó aleatoriamente a uno de los 10 conjuntos, la siguiente a otro conjunto diferente, y así sucesivamente. Se continuó asignando las palabras con mejor puntuación en bloques de 10 hasta completar los 10 conjuntos. Este método buscó una distribución equitativa de las palabras con mejor reconocimiento en todas las listas que debió validarse posteriormente comparando las curvas psicométricas. Una vez generadas las 10 listas, se tomaron los resultados de reconocimiento para cada nivel de presentación, para cada lista. Por ejemplo, se sumaron los resultados de reconocimiento de cada una de las palabras de la lista 1 para el nivel -5 dB y se calculó el porcentaje de aciertos respecto a las 25 palabras de la lista. Repitiendo este procedimiento para los restantes niveles, se obtuvo la curva psicométrica para la lista 1. Lo mismo se hizo para el resto de las listas. Una vez generado el conjunto, se verificó el balance fonético de cada una de las listas. De ser necesario, para mejorar dicho balance, se intercambiaron algunas palabras con puntaje similar entre distintas listas.

Para obtener la función psicométrica de cada lista, se realizó un ajuste a una función sigmoide sobre las respuestas binarias (es decir, correcto o incorrecto) de los sujetos con la siguiente función logística:

$$p(x; m, w) = \frac{1}{1 + e^{-2 \log\left(\frac{1}{.05-1}\right) \frac{x-m}{w}}}$$

La curva psicométrica se estimó para cada lista utilizando la herramienta de MATLAB Psignifit Toolbox⁽¹⁸⁾.

Tabla 1. Listas de palabras definitivas para pruebas de porcentaje de reconocimiento de palabras

Lista 1	Lista 2	Lista 3	Lista 4	Lista 5	Lista 6	Lista 7	Lista 8	Lista 9	Lista 10
Perro	Verde	Hora	Norte	Algo	Chico	Ropa	Nada	Cerca	Aire
Sola	Tiempo	Cinco	Árbol	Clara	Hola	Grande	Mezcla	Uno	Creo
Triste	Donde	Tarde	Oro	Leche	Súper	Flores	Viven	Frío	Cama
Loco	Fría	Chica	Miedo	Piso	Siete	Dijo	Río	Claro	Gente
Cara	Disco	Clave	Cuan-do	Centro	Como	Casa	Calle	Lejos	Una
Alguien	Solo	Orden	Quiero	Forma	Caja	Tema	Dicen	Entre	Suelo
Traje	Cuales	Todo	Para	Otra	Nadie	Luna	Julio	Pico	Plata
Vamos	Pobre	Nues-tra	Tía	Pelo	Libro	Bosque	Frente	Diario	Hasta
Mio	Hizo	Mapa	Meses	Día	Ella	Tienen	Alto	Dicho	Nom-bre
Nunca	Cola	Era	Linda	Justo	Eso	Arte	Cosas	Habla	Medio
Pero	Cero	Dulce	Pudo	Quince	Ruta	Toda	Pueden	Once	Negro
Simple	Mía	Cine	Jefe	Cielo	Trece	Porque	Sigo	Lunes	Tío
Manos	Mente	Dije	Virus	Martes	Fácil	Siguen	Cierto	Menos	Quie-ran
Dado	Tuvo	Siglo	Campo	Ritmo	Lindo	Siem-pre	Poco	Contra	Antes
Hacia	Rico	Viste	Alta	Viernes	Tengan	Carne	Letra	Tuve	Redes
Desde	Mira	Quinto	Aunque	Blanco	Niña	Cuerpo	Grupo	Blanca	Queda
Hablo	Cada	Sala	Seres	Llevan	Clases	Hablan	Hom-bre	Serie	Malo
Dice	Hacen	Parque	Estos	Pena	Ponen	Libre	Hice	Llega	Crisis
Miren	Salen	Dentro	Mundo	Hace	Nues-tros	Quede	Mala	Muerte	Fondo
Cree	Suena	Esas	Sigue	Dudas	Hemos	Estas	Anda	Diga	Casi
Punto	Digo	Vemos	Parte	Mire	Marca	Miles	Ido	Saca	Ese
Cuenta	Tipo	Llaman	Caso	Deben	Vino	Uso	Mesa	Misma	Listo
Sirve	Entra	Siento	Lista	Cuen-tos	Doble	Niño	Paso	Vimos	Libros
Bloque	Plaza	Cumple	Liena	Marco	Corte	Mismos	Cuanto	Puede	Hubo
Sean	Quie-nes	Verlo	Tiene	Pasa	Darle	Lados	Este	Sino	Planta

Resultados

El conjunto definitivo de listas, 10 listas con 25 palabras cada una, obtenidas a partir del proceso de puntuación y balance fonético, se muestran en la Tabla 1.

Para el trazado de las curvas psicométricas, se utilizó un *software* desarrollado en MATLAB que, en base a los datos registrados, realizó el ajuste de estos mediante regresión logística. En la Figura 1 se muestran las funciones psicométricas de cada conjunto de listas calculadas mediante regresión logística.

En la Tabla 2, pueden observarse los valores de la pendiente de cada función (derivada en el punto de 50% de aciertos) y la pendiente calculada en base al rango entre el 20% y 80%.

Se observa en la Tabla 2 un desvío estándar de 0.36 para la pendiente en el 50% y 0.23 para el 20%-80%.

También se verificó la diferencia de dB HL entre listas para el umbral del 50%. En este caso, el desvío

Figura 1. Curvas psicométricas de las 10 listas de palabras, sin corrección de nivel

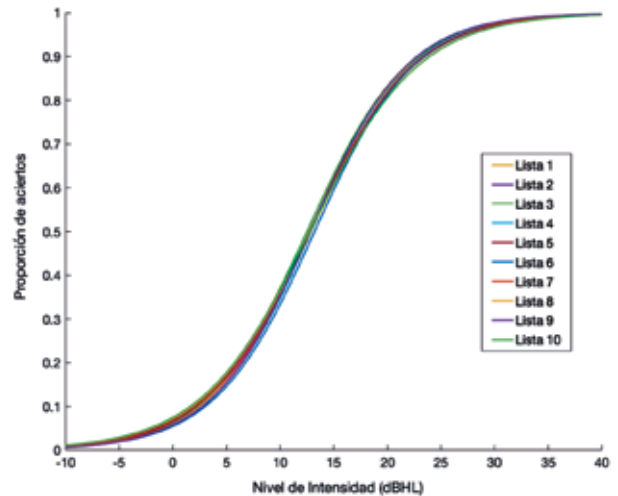


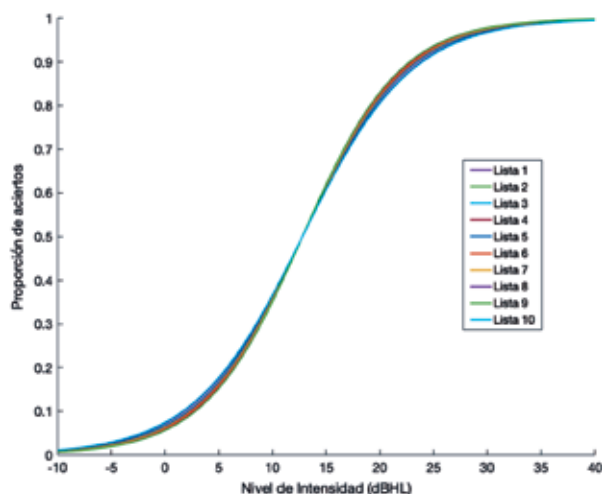
Tabla 2. Parámetros de las curvas psicométricas obtenidas para las listas conformadas

Lista	Umbral	Pendiente 50%	Pendiente 20%-80%	Δ dBHL
	dB HL	%/dB	%/dB	
1	12.17	5.65	3.61	-0.17
2	12.42	5.27	3.37	-0.42
3	12.3	5.46	3.49	-0.3
4	12.36	5.29	3.39	-0.36
5	12.06	5.78	3.7	-0.06
6	12.23	6.19	3.96	-0.23
7	12.22	5.4	3.46	-0.22
8	12.27	5.44	3.48	-0.27
9	12.33	6.16	3.94	-0.33
10	11.86	6.13	3.92	0.14
Media	12.22	5.68	3.63	-0.22
Mínimo	11.86	5.27	3.37	0.14
Máximo	12.42	6.19	3.96	-0.42
Desv. Est.	0.16	0.36	0.23	

estándar es de 0.16 dB y un valor medio de 12.22 dB. Para lograr que todas las curvas compartan el mismo umbral del 50%, se estableció un valor de referencia de 12 dB HL y se modificaron digitalmente en nivel los archivos correspondientes a cada lista con la diferencia calculada. De esta forma, todas las curvas presentan el 50% de aciertos en el nivel de intensidad de 12 dB HL. Las curvas resultantes se muestran en la Figura 2 (en pág. siguiente).

Las pendientes medidas, como la derivada en el punto correspondiente al 50% de aciertos de las funciones psicométricas encontradas en este estudio (5.68%/dB en promedio), son similares a los valores previamente reportados para distintos materiales de reconocimiento de palabras desarrollados en inglés y otros idiomas^(4, 5, 9).

Figura 2. Curvas psicométricas de las 10 listas de palabras luego del ajuste de nivel para que crucen el 50% de aciertos en 12 dB HL.



Discusión

Tradicionalmente, se asignó gran importancia al balance fonético del material de cada lista. Se entiende por balance fonético a que la proporción de los distintos fonemas dentro de la lista respete la proporción de aparición de los fonemas en el idioma considerado. Es importante aclarar, tal como mencionan Lehiste y Peterson⁽¹⁹⁾, que el balance fonético implica una representación equitativa de sonidos del habla, aunque este tipo de balance es limitado debido a las interacciones entre distintos sonidos. Sería más apropiado y posible lograr el balance fonémico, donde los fonemas están representados de manera equilibrada. Es más común, por costumbre, hablar de balance fonético, haciendo esta salvedad. La importancia del balance fonético en la generación de este tipo de material fue, sin embargo, puesta en duda por un trabajo de Martin et al.⁽²⁰⁾, quienes encontraron que el porcentaje de reconocimiento de palabras no variaba de manera significativa utilizando un material balanceado o un conjunto aleatorio de palabras. El estudio de Martin incluyó dos grupos de sujetos: uno con audición normal y otro con pérdidas auditivas neurosensoriales. Sin embargo, no se proporciona información detallada sobre el perfil audiométrico específico de los sujetos con pérdidas auditivas, como la gravedad, tipo y configuración de la pérdida. Un pronunciado desbalance fonético puede impactar de manera diferente en distintas configuraciones de pérdida auditiva. Por eso se considera que un grado de balance fonético puede ser aconsejable, como criterio ordenador. En este trabajo, si bien se ha adoptado el balance foné-

tico como un criterio a respetar, se priorizó la equivalencia psicométrica entre las listas.

Balance fonético

Se estudió el balance fonético del material del Dr. Tato et al. y del material propuesto en el presente trabajo. Para ello se registraron las frecuencias de aparición de cada fonema en cada una de las listas. Por otra parte, se registró la frecuencia de aparición de los diferentes fonemas en el conjunto de 23 millones de palabras recopiladas. Los resultados se muestran en la Tabla 3 (en pág. siguiente), donde pueden observarse las frecuencias de aparición de los diferentes fonemas.

Una forma directa y simple para medir la similitud entre dos distribuciones de frecuencias podría ser el cálculo del coeficiente de correlación de Pearson. Coloma⁽²¹⁾ utilizó y desestimó esta técnica para analizar el balance fonético. En su trabajo señaló que, aun obteniendo valores altos de correlación este tipo de cálculo, por su naturaleza, promedia las diferencias y puede esconder algunos efectos en algunos fonemas individuales. Por ello propuso el uso de otras herramientas, basadas en el análisis de concordancia de Bland-Altman. En otro trabajo, Aubanel et al.⁽²²⁾ utilizaron el cuadrado de la distancia euclidiana entre las frecuencias esperadas (del idioma) y las frecuencias presentes en el corpus bajo análisis, haciendo uso de la fórmula:

$$d_{sc} = \sum_{p=1}^P (f_{ps} - f_{pc})^2$$

Donde f_{ps} y f_{pc} son las frecuencias de aparición de cada fonema en el idioma y en el conjunto considerado. Cuanto mayor es este número, menor será el grado de balance fonémico.

En este trabajo se comparó el balance fonético entre las listas de Tato et al. y las nuevas listas propuestas. Para ello, se calcularon los coeficientes de correlación de Pearson (R) para las listas de Tato et al. y las nuevas listas. También se calculó la distancia euclídea para ambos materiales y se efectuó un análisis de concordancia según el método de Bland-Altman. Los resultados se muestran en la Tabla 4 (en pág. siguiente).

Estos datos muestran el excelente balance fonético que presentan las listas de Tato et al., siendo la media de los coeficientes de correlación R de las listas de 0.9814. Para las nuevas listas, la media del valor R es aún más elevado, resultando igual a 0.98883. La distancia euclídea al cuadrado pro-

Tabla 3. Frecuencia de aparición de los fonemas en español rioplatense para el conjunto total (23 millones de palabras), las listas de Tato et al. y las listas propuestas

ER	Tato A1	Tato A2	Tato A3	Tato A4	Tato B1	Tato B2	Tato B3	Tato B4	Tato C1	Tato C2	Tato C3	Tato C4	Nueva1	Nueva2	Nueva3	Nueva4	Nueva5	Nueva6	Nueva7	Nueva8	Nueva9	Nueva10	
ɲ	0,16	0,00	0,00	0,00	0,87	0,00	0,87	0,00	0,87	0,00	0,86	0,00	0,87	0,00	0,00	0,00	0,00	0,84	0,86	0,00	0,00	0,00	
ʃ	0,29	0,00	0,00	0,00	0,87	0,00	0,00	0,00	0,87	0,00	0,00	0,00	0,87	0,00	0,00	0,83	0,00	0,83	0,84	0,00	0,00	0,87	0,87
ç	0,30	0,00	0,00	0,00	0,00	0,87	0,87	0,00	0,00	0,86	0,00	0,00	0,87	0,00	0,00	0,00	0,00	0,84	0,86	0,88	0,00	0,00	0,00
x	0,33	0,00	0,87	0,00	0,87	0,00	0,00	0,86	0,87	0,00	0,86	0,87	0,87	0,00	0,00	0,00	0,00	0,83	0,84	0,86	0,88	0,87	0,00
ç	0,34	0,00	0,00	0,87	0,00	0,00	0,87	0,00	0,00	0,00	0,00	0,00	0,86	0,00	0,83	0,85	0,00	0,00	0,00	0,00	0,00	0,00	0,87
ŋ	0,44	0,86	0,87	0,00	0,00	0,87	0,87	0,00	0,00	0,00	0,00	0,00	0,86	0,00	0,83	0,85	0,83	0,84	0,00	0,00	0,87	0,00	0,00
f	0,49	0,86	0,87	0,00	0,87	0,87	0,00	0,00	0,87	0,86	0,00	0,00	0,87	0,00	0,00	0,83	0,85	0,83	0,84	0,00	0,88	0,87	0,00
r	0,51	0,86	0,87	0,87	0,87	1,74	0,87	0,86	0,87	0,86	0,86	0,87	0,87	0,86	0,88	0,00	0,85	0,83	0,84	0,86	0,88	0,00	0,87
ʃ	0,59	0,86	0,00	0,00	0,00	0,87	0,87	0,86	0,00	0,86	0,00	0,87	0,00	0,00	0,88	0,00	0,85	0,83	0,84	0,86	0,88	0,87	0,87
y	0,66	0,86	0,87	0,87	0,87	0,00	0,00	0,86	0,87	0,00	0,86	0,87	0,00	1,72	0,88	0,83	0,85	0,83	0,00	0,86	0,00	1,74	0,87
w	1,12	0,00	0,87	0,00	0,87	0,87	0,87	0,86	0,87	0,00	0,86	1,74	0,87	0,86	1,77	1,67	1,69	0,83	0,84	0,86	1,75	0,87	1,74
β	1,26	0,86	1,74	0,87	0,00	1,74	0,87	0,86	0,00	1,72	0,86	0,87	0,87	0,86	1,77	0,83	1,69	1,65	1,68	1,72	0,88	1,74	0,87
b	1,29	1,72	0,87	1,74	1,74	1,74	1,74	1,72	1,74	0,86	1,72	1,74	0,87	1,72	0,88	1,67	0,85	1,65	0,84	0,86	1,75	1,74	0,87
h	1,50	0,86	1,74	1,74	0,87	0,87	0,00	0,00	1,74	0,00	0,86	0,87	0,87	1,72	0,88	1,67	1,69	2,48	0,84	2,59	1,75	0,87	1,74
j	1,55	1,72	2,61	0,87	2,61	1,74	2,61	2,59	0,87	2,59	1,72	1,74	2,61	1,72	1,77	0,83	1,69	1,65	1,68	0,86	1,75	2,61	2,61
ð	1,77	0,00	0,87	3,48	1,74	2,61	2,61	2,59	1,74	3,45	1,72	2,61	2,61	1,72	1,77	2,50	2,54	0,83	1,68	2,59	2,63	0,87	2,61
u	2,05	3,45	1,74	3,48	2,61	2,61	1,74	2,59	2,61	2,59	1,72	1,74	2,61	1,72	0,88	1,67	2,54	1,65	2,52	1,72	1,75	2,61	0,87
d	2,55	4,31	3,48	1,74	3,48	1,74	1,74	2,59	3,48	0,86	2,59	2,61	2,61	3,45	3,54	2,50	2,54	2,48	1,68	1,72	1,75	2,61	1,74
p	2,79	2,59	2,61	3,48	2,61	2,61	2,61	2,59	3,48	2,59	4,31	2,61	2,61	2,59	3,54	3,33	2,54	2,48	3,36	2,59	2,63	1,74	1,74
m	3,42	2,59	2,61	2,61	2,61	2,61	2,61	2,59	2,61	2,59	2,61	2,61	3,45	2,74	3,33	4,24	4,13	3,36	4,31	4,39	4,35	4,35	4,35
l	4,52	5,17	5,22	5,22	5,22	5,22	5,22	5,17	5,22	4,31	4,31	5,22	5,22	5,17	4,42	5,83	5,08	4,96	5,04	5,17	5,26	5,22	5,22
k	4,64	5,17	2,61	3,48	3,48	4,35	3,48	3,45	3,48	5,17	3,45	3,48	3,48	5,17	4,42	4,17	4,24	4,96	5,04	5,17	4,39	4,35	4,35
t	4,73	4,31	4,35	4,35	4,35	4,35	4,35	5,17	3,48	4,31	4,31	4,35	4,35	5,17	5,31	5,83	4,24	4,96	5,04	4,31	5,26	3,48	5,22
i	5,59	6,03	5,22	6,09	4,35	6,09	4,35	4,31	6,09	6,03	6,03	5,22	4,35	5,17	6,19	5,83	5,08	4,96	5,04	6,03	6,14	5,22	5,22
r	5,64	5,17	6,09	6,09	5,22	4,35	6,09	6,03	5,22	5,17	6,03	6,09	5,22	6,03	5,31	6,67	5,93	5,79	5,04	6,03	5,26	6,09	6,09
n	6,79	6,90	6,96	7,83	7,83	6,96	6,96	7,76	7,83	7,76	7,76	8,70	7,83	6,03	7,08	5,83	5,93	5,79	5,88	6,03	6,14	6,09	6,09
s	7,87	9,48	8,70	7,83	8,70	8,70	9,57	9,48	7,83	11,21	9,48	7,83	8,70	6,90	7,96	7,50	7,63	7,41	8,40	6,90	7,88	7,83	6,96
o	10,16	8,62	9,57	10,43	9,57	8,70	10,43	9,48	9,57	8,62	9,48	9,57	10,34	10,63	9,17	10,17	10,74	9,21	10,34	10,53	9,57	11,30	11,30
a	12,42	12,93	13,04	13,04	13,91	13,04	12,17	12,93	13,91	12,93	13,04	13,91	12,93	13,27	11,67	11,02	11,60	11,76	11,21	11,38	13,04	13,04	13,04
e	14,21	13,79	13,91	13,04	13,04	13,91	14,78	13,79	13,04	13,79	13,91	13,04	12,93	13,26	13,33	13,56	13,22	13,41	13,79	13,16	13,04	13,04	13,04

Tabla 4. Coeficiente de correlación R, cuadrado de la distancia euclídea, media de las diferencias e intervalo de confianza de las diferencias al 95% para las listas de Tato et al. y las nuevas listas

LISTA	R(PEARSON)	DIST. EUCL%	DIF. MEDIA%	CI (95%)
Tato A1	0,9779	18,6160	-0,0001	3,1407
Tato A2	0,9853	12,1935	0,0289	2,5392
Tato A3	0,9814	15,7541	-0,0001	2,8892
Tato A4	0,9796	16,8082	-0,0001	2,9843
Tato B1	0,9848	12,2649	-0,0001	2,5493
Tato B2	0,9825	15,3887	-0,0001	2,8555
Tato B3	0,9826	14,9474	-0,0001	2,8143
Tato B4	0,9809	15,7111	-0,0001	2,8853
Tato C1	0,9680	29,0718	-0,0001	3,9248
Tato C2	0,9858	12,4106	-0,0001	2,5644
Tato C3	0,9874	10,7250	-0,0001	2,3839
Tato C4	0,9807	15,8315	-0,0001	2,8963
MEDIA TATO	0,9814	15,8102	0,0023	2,8690
Nueva1	0,9893	8,7328	-0,0001	2,1511
Nueva2	0,9901	8,5927	-0,0026	2,1338
Nueva3	0,9863	11,4887	-0,0001	2,4673
Nueva4	0,9924	7,6807	-0,0001	2,0174
Nueva5	0,9908	8,1817	-0,0002	2,0821
Nueva6	0,9906	8,8501	0,0305	2,1621
Nueva7	0,9870	10,8292	-0,0001	2,3954
Nueva8	0,9892	9,2358	-0,0281	2,2094
Nueva9	0,9843	12,8485	-0,0001	2,6092
Nueva10	0,9832	13,6748	-0,0001	2,6918
MEDIA NUEVAS	0,9883	10,0115	-0,0001	2,2920

medio arrojó un valor medio de 15.8 para las listas de Tato et al. y un valor medio de 10 para las nuevas. Un valor más bajo de este parámetro im-

plica un mayor balance fonémico. Por último, del análisis de Bland-Altman, se extraen los valores de la media de las diferencias y el ancho del intervalo de confianza al 95% (calculado mediante la expresión $2x1,96xSD$). Todos los indicadores comparados arrojan resultados favorables a las nuevas listas en cuanto a balance fonémico.

Equivalencia psicométrica entre listas

La equivalencia psicométrica de las diferentes listas es una condición importante para su aplicación clínica. Uno de los usos frecuentes de este tipo de material, junto con la evaluación que forma parte de la logaudiometría, es la aplicación en el estudio de validación del beneficio de un equipamiento auditivo. Es importante que las listas presenten equivalencia psicométrica para poder comparar efectivamente distintos dispositivos de ayuda auditiva con distintas listas.

El propósito de este estudio fue crear un conjunto de listas de palabras bisilábicas de acentuación grave en español rioplatense, relativamente homogéneas para su uso en la medición del reconocimiento de palabras. La inspección visual de la Figura 3 indica que, para sujetos con audición normal, las listas son homogéneas con respecto a la audibilidad y la pendiente de la función psicométrica, especialmente

después del ajuste digital. Se realizó una prueba de Kruskal-Wallis para muestras independientes para evaluar si existen diferencias significativas en las medianas entre las 10 listas. El estadístico de prueba H fue de 0.143 con 9 grados de libertad ($p > 0.05$), por lo tanto, no se rechaza la hipótesis nula. Esto sugiere que no hay evidencia suficiente para afirmar que existen diferencias significativas entre las listas.

Conclusión

La presentación de este conjunto de listas de palabras grabado para obtener el WRS en español rioplatense tiene por objeto solucionar los problemas observados en las listas del Dr. Tato et al. Contienen palabras de uso actual y pueden resultar útiles para uso clínico y en investigación. Las 10 listas de 25 palabras bisilábicas graves presentan equivalencia psicométrica. Las pendientes de las curvas psicométricas que corresponden a las listas se encuentran dentro de los rangos habituales para este tipo de material. El balance fonético del material presenta valores similares, o mejores, a las listas utilizadas hasta el presente, y la equivalencia psicométrica entre las listas fue verificada. En forma conjunta con el material específico desarrollado para obtener el SRT, este material puede servir como base para la realización de estudios logoaudiométricos. El material puede ser utilizado para evaluar el rendimiento comparativo de prótesis auditivas en la fase de validación y en general cuando sea necesario conocer la capacidad de reconocimiento de palabras en español rioplatense. El material grabado fue validado en este estudio y se encuentra disponible para descarga en: <https://cistas.untref.edu.ar/academia>

Agradecimiento

Los autores agradecen al Ing. Mariano Girola y a la Fga. Silvia Mastroianni por sus valiosos aportes a este trabajo.

Los autores no manifiestan conflictos de interés.

Bibliografía

- Katz Jack, Chasin Marshall, English KM., Hood LJ., Tillery KL. *Handbook of clinical audiology* (7th ed.). Philadelphia, PA: Wolters Kluwer. 2015:61-72.
- Cristiani H, Piegari A, Alonso S, Virgallito O, Ausili S. Desarrollo de un conjunto de palabras psicométricamente equivalentes para obtener el umbral de reconocimiento del habla en Español Rioplatense. *Federación Argentina de Sociedades de Otorrinolaringología*. 2024;1:19-27.
- Carlo MA, Wilson RH, Villanueva-Reyes A. Psychometric Characteristics of Spanish Monosyllabic, Bisyllabic, and Trisyllabic Words for Use in Word-Recognition Protocols. *J Am Acad Audiol*. el 1 de julio de 2020;31(7):531-46.
- Harris RW, Nissen SL, Pola MG, McPherson DL, Tavartkiladze GA, Eggett DL. Psychometrically equivalent Russian speech audiometry materials by male and female talkers. *Int J Audiol*. enero de 2007;46(1):47-66.
- Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development Of Materials For Speech Audiometry. *Journal of Speech and Hearing Disorders*. septiembre de 1952;17(3):321-37.
- TILLMAN TW, Carhart R. An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory test No. 6. 1966 jun.
- Martin Michael. *Speech audiometry*. Taylor & Francis; 1987, 327-328.
- Dukes AJ. Psychometrically Equivalent Bisyllabic Word Lists for Word Psychometrically Equivalent Bisyllabic Word Lists for Word Recognition Testing in Taiwan Mandarin Recognition Testing in Taiwan Mandarin. 2006 [citado el 24 de septiembre de 2024]; Disponible en: <https://scholarsarchive.byu.edu/etd/460>
- Trimmis N, Papadeas E, Papadas T, Naxakis S, Papatathanasopoulos P, Goumas P. *Speech Audiometry: The Development of Modern Greek Word Lists for Suprathreshold Word Recognition Testing*. *The Mediterranean Journal of Otolaryngology*. 2005: 117-126.
- Tato JM, Lorente Sanjurjo F, Bello J. Características acústicas de nuestro idioma. *Federación Argentina de Sociedades de Otorrinolaringología*. 2004: 67-72.
- Sala L. Una nueva mirada sobre las listas de palabras fonéticamente balanceadas. [Argentina]: Universidad Faosta; 2012.
- Pace M. Desarrollo de nuevas listas de palabras psicométricamente equivalentes y fonéticamente balanceadas, para su uso en logoaudiometrías y otras pruebas de reconocimiento del habla con español rioplatense. [Argentina]: UNTREF; 2023.
- Mortensen DR, Dalmia S, Littell P. *Epitran: Precision G2P for Many Languages*. *International Conference on Language Resources and Evaluation*. 2018: 2710-2714.
- Arias Rodríguez I. Frequency of occurrence of phonemes and allophones in contemporary Spanish as calculated by an automatic transcription system. *Loquens*. 2016;3(1) e029.
- Hernández-Figueroa Z, Carreras-Riudavets FJ, Rodríguez-Rodríguez G. Automatic syllabification for Spanish using lemmatization and derivation to solve the prefix's prominence issue. *Expert Syst Appl*. 2013;40(17):22-31.
- ANSI/ASA S3.6-2010 - Specification for Audiometers.
- ANSI/ASA S3.1-1999 (R2018) - Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms.
- Schütt HH, Harmeling S, Macke JH, Wichmann FA. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Res*. mayo de 2016;122:105-123.
- Lehiste I, Peterson GE. Linguistic Considerations in the Study of Speech Intelligibility. *J Acoust Soc Am*. el 1 de marzo de 1959;31(3):280-286.
- Martin FN, Champlin CA, Perez DD. The question of phonetic balance in word recognition testing. *J Am Acad Audiol*. octubre de 2000;11(9):489-493.
- Coloma G. A statistical comparison between two texts to illustrate the phonetics of Spanish. CEMA Working Papers: Serie Documentos de Trabajo [Internet]. 2017 [citado el 24 de septiembre de 2024]; Disponible en: <https://ideas.repec.org/p/cem/doctra/619.html>
- Aubanel V, Lecumberri MLG, Cooke M. The Sharvard Corpus: A phonemically-balanced Spanish sentence resource for audiology. *Int J Audiol*. el 26 de septiembre de 2014;53(9):633-638.