

RECONOCIMIENTO DE LA PALABRA BAJO UMBRAL ÓPTIMO SIMPLIFICADO: REVISIÓN DE ALCANCE DE LISTAS DE PALABRAS ORALES VS. DIGITALES

Word recognition under simplified optimal presentation level: a scoping review of oral vs. digital word lists

GÉNESIS OLGUÍN HURTADO 

Universidad del Museo Social Argentino. Doctorado en Fonoaudiología, Argentina

MARIANO GUILLERMO BLAKE 

Universidad de Buenos Aires - CONICET. Instituto de Fisiología y Biofísica

Bernardo Houssay. Facultad de Medicina. Buenos Aires, Argentina.

Universidad del Museo Social Argentino. Laboratorio de Investigaciones

Fonoaudiológicas. Facultad de Ciencias Humanas. Buenos Aires, Argentina.

Recibido: 01/09/2025

Aceptado: 28/10/2025

Autor de correspondencia: Av. Corrientes 1723 (C1042AAD), Ciudad Autónoma de Buenos Aires, Argentina.
Correo electrónico: fono.genesisol.hurtado@gmail.com

RESUMEN

La evaluación del reconocimiento de palabras es fundamental en la práctica audiológica contemporánea, donde existe un debate creciente sobre las ventajas de las pruebas digitalizadas frente a la presentación oral tradicional. Esta revisión de alcance analizó estudios publicados entre 2020 y 2025, siguiendo la metodología PRISMA-ScR, con el objetivo de comparar ambas modalidades y examinar la aplicación del umbral óptimo simplificado (PTP + 30 dB).

Los hallazgos evidencian que las listas digitalizadas ofrecen un mayor grado de confiabilidad y consistencia que la presentación en vivo. Se identificó un desarrollo notable de herramientas tecnológicas, incluidas aplicaciones móviles, que favorecen la automatización y estandarización de las evaluaciones. No obstante, se observaron limitaciones relevantes, como la baja adopción del criterio de intensidad fija (PTP + 30 dB) y la ausencia generalizada de análisis estadísticos robustos, por ejemplo, los métodos de Bland-Altman o el coeficiente de correlación intraclass.

Aunque la digitalización presenta fundamentos metodológicos sólidos, se requieren investigaciones que incorporen diseños comparativos rigurosos, análisis avanzados de confiabilidad y validación mediante correlatos neurofisiológicos, con el fin de consolidar su implementación como estándar clínico.

Palabras clave: Aplicaciones Móviles; Revisiones Sistemáticas como Asunto; Audiología; Software de Reconocimiento del Habla; Pruebas de Asociación de Palabras; Semántica.



1. Introducción

El reconocimiento de la palabra es una de las pruebas más relevantes en la evaluación audiológica. A diferencia de la audiometría tonal, que mide únicamente la detección de sonidos puros, esta prueba evalúa la capacidad funcional de comprender el habla en situaciones cotidianas. Esta distinción no es trivial, pues la habilidad para reconocer palabras influye directamente en la participación social, el desempeño académico, las oportunidades laborales y la calidad de vida de las personas con pérdida auditiva [1]. Los resultados de estas evaluaciones orientan decisiones clínicas trascendentes, como por ejemplo en el ajuste de audífonos, indicación de implantes cocleares, evaluación del beneficio post-intervención o definición de estrategias de rehabilitación. Incluso pueden determinar la idoneidad laboral en contextos específicos. Por tanto, la validez y confiabilidad de estas pruebas no constituyen aspectos meramente metodológicos, sino factores que impactan directamente en la vida de las personas y en la eficiencia de los programas de salud auditiva.

Durante décadas, la presentación oral fue el método estándar. El evaluador pronunciaba las palabras directamente frente al paciente, ajustando la intensidad de forma intuitiva y supervisando las respuestas en tiempo real. Este enfoque resultó práctico y accesible, pero gradualmente evidenció limitaciones importantes. Variables como la articulación, la entonación, la velocidad de dictado, la fatiga vocal o alteraciones fonatorias del examinador introducían inconsistencias que dificultaban la comparación entre profesionales, sesiones o centros clínicos [2,3,4]. Esto generaba incertidumbre y cuestionamientos, que hacían difícil establecer si las diferencias obtenidas en los puntajes reflejaban cambios auditivos reales en la persona o simplemente variaciones en la aplicación de la prueba.

Frente a estas inconsistencias, la digitalización se ha consolidado como una alternativa metodológica sólida. Las grabaciones controladas y calibradas acústicamente aseguran que cada estímulo se presente con intensidad constante, eliminando la variabilidad inherente al factor humano. Adicionalmente, esta modalidad permite aleatorizar estímulos sistemáticamente, reduciendo efectos de memoria, automatizar el registro de respuestas y mejorar la precisión del análisis. Estas características incrementan la reproducibilidad y facilitan comparaciones confiables entre diferentes contextos clínicos e investigativos [5]. La digitalización también ha impulsado la teleaudiología, especialmente en zonas rurales o con escasez de especialistas. En estos entornos, las listas estandarizadas garantizan que la evaluación mantenga su calidad independientemente del lugar de aplicación. Desde la perspectiva de salud pública, esto contribuye a reducir inequidades en el acceso y ampliar el alcance de los programas de detección y rehabilitación auditiva [2, 4, 5].

Sin embargo, persiste un debate metodológico fundamental acerca de la intensidad a la cual deben presentarse los estímulos. Idealmente, se debería obtener el umbral de máxima discriminación (UMD) mediante curvas psicométricas completas, pero este procedimiento demanda tiempo, personal capacitado y recursos que no siempre están disponibles. Una alternativa es el umbral óptimo simplificado, que fija la intensidad en 30 dB por encima del promedio tonal puro (PTP + 30 dB). Este criterio busca reducir la variabilidad asociada a ajustes individuales y homogeneizar procedimientos, especialmente en centros con recursos limitados [6]. Debido a que en la literatura en español este criterio suele denominarse umbral de máxima discriminación, en adelante se empleará la sigla UMD para referirse a ello.

No obstante, la literatura reciente revela que este enfoque ha sido escasamente explorado. Si bien la mayoría de los estudios reconocen las ventajas de las listas digitalizadas, pocos han incorporado el criterio PTP + 30 dB de manera sistemática. La validación estadística también permanece en un nivel básico, debido a que predominan pruebas como t de Student o Wilcoxon, mientras que herramientas más robustas, como el coeficiente de correlación intraclass, los modelos mixtos o el análisis de Bland-

Altman continúan sin utilizarse [7, 8]. Tampoco se han integrado mediciones neurofisiológicas objetivas, como potenciales evocados auditivos del tronco encefálico (PEATC) o respuestas auditivas de estado estable (ASSR), que podrían confirmar hallazgos importantes y profundizar en los mecanismos del procesamiento del habla.

A estas limitaciones se añade la dimensión transcultural. Aunque algunos estudios han digitalizado listas en español, portugués o árabe, las diferencias prosódicas y fonéticas entre lenguas exigen procesos rigurosos de adaptación y validación. Sin análisis comparativos que aseguren la equivalencia de los estímulos, resulta difícil establecer protocolos globalmente aplicables.

En este contexto de avances y vacíos, se hace necesaria una síntesis sistematizada que ordene el conocimiento disponible. Este trabajo ofrece la primera revisión de alcance sobre la aplicación del umbral óptimo simplificado en pruebas digitalizadas de reconocimiento de la palabra, analizando estudios publicados entre 2020 y 2025. El análisis se centra en tres ejes: la comparación entre modalidades oral y digital, la viabilidad del criterio PTP + 30 dB y las innovaciones tecnológicas asociadas a la digitalización de listas de palabras.

2. Material y Métodos

Esta revisión se desarrolló siguiendo las directrices PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) y las etapas metodológicas propuestas por Arksey y O’Malley: identificación del problema, búsqueda de literatura, selección de estudios, extracción de datos y síntesis de información [9,10]. Esta estrategia permitió representar de manera comprensiva el estado del conocimiento en un campo con fundamentos teóricos consolidados, pero con notables variaciones en su aplicación práctica y desarrollo tecnológico.

Estrategia de búsqueda

Se realizó una búsqueda sistemática entre marzo y mayo de 2025 en seis bases de datos electrónicas: PubMed/MEDLINE, Scopus, Web of Science (WOS), SciELO, LILACS y Google Scholar. No se aplicaron restricciones por tipo de estudio. La búsqueda se orientó a identificar evidencia relacionada con el reconocimiento de palabras, el umbral óptimo simplificado y las modalidades de presentación de estímulos verbales (oral frente a digital) en audiología clínica.

Se incluyeron artículos publicados entre 2020 y 2025, en inglés, español y portugués. Este período se justifica por dos razones principales. Primero, coincide con un notable crecimiento en el desarrollo y validación de herramientas digitales en audiología, impulsado parcialmente por la necesidad de adaptar las evaluaciones clínicas durante la pandemia de COVID-19. Segundo, permite recopilar evidencia reciente sobre la implementación del criterio PTP + 30 dB y su integración con tecnologías digitales, reflejando las tendencias actuales en práctica clínica e investigación audiológica.

La formulación base de búsqueda fue:

(“speech audiometry” OR “word recognition” OR “audiometría verbal” OR “reconocimiento de palabras” OR “word list” OR “lista de palabras”) AND (“maximum discrimination” OR “word recognition score” OR WRS OR PBmax OR “umbral de máxima discriminación”) AND (“live voice” OR “recorded voice” OR “digital recording” OR “voz viva” OR “grabación digital”)

Esta formulación fue adaptada según las especificaciones de cada base de datos:

- PubMed (n = 42): Campos [tiab], filtro 2020–2025, inglés. Cadena: (“word recognition”[tiab] OR “speech audiometry”[tiab] OR “word list”[tiab] OR “maximum discrimination”[tiab] OR “word recognition score”[tiab] OR WRS[tiab] OR PBmax[tiab] OR “live voice”[tiab] OR “recorded voice”[tiab] OR “digital recording”[tiab]).
- Scopus (n = 29): Campos TITLE-ABS-KEY, filtro 2020–2025, inglés/español/portugués. Cadena: TITLE-ABS-KEY(“word recognition” OR “speech audiometry” OR “word list” OR “audiology” OR “maximum discrimination” OR “word recognition score” OR WRS OR PBmax OR “live voice” OR “recorded voice” OR “digital recording”).
- Web of Science (n = 20): Campo TS (Topic), filtro 2020–2025, sin restricción idiomática. Cadena: TS=(“word recognition” OR “speech audiometry” OR “word list” OR “audiology” OR “maximum discrimination” OR “word recognition score” OR WRS OR PBmax OR “live voice” OR “recorded voice” OR “digital recording”).
- SciELO (n = 18): Texto completo, filtro 2020–2025, español y portugués.
 - Español: (“reconocimiento de palabras” OR “audiometría verbal” OR “lista de palabras” OR “umbral de máxima discriminación” OR “grabación digital” OR “presentación oral”).
 - Portugués: (“reconhecimento de palavras” OR “audiometria vocal” OR “lista de palavras” OR “discriminação máxima” OR “gravação digital” OR “voz ao vivo”).
- LILACS (n = 15): Portal BVS, filtro 2020–2025, español y portugués. Cadena: (“reconocimiento de palabras” OR “audiometría verbal” OR “lista de palabras” OR “umbral de máxima discriminación” OR “grabación digital” OR “presentación oral”).
- Google Scholar (n = 8): Texto completo, filtro 2020–2025. Cadena: “word recognition score” “digital recording” “live voice” “word list”.

Se reconoce que el número de resultados puede variar entre bases de datos debido a diferencias en cobertura temática, estructura de indexación, compatibilidad con operadores booleanos y presencia de literatura gris.

La selección y extracción de artículos fue realizada de forma independiente y a doble ciego por dos investigadores, utilizando el software Rayyan® para organizar los registros y aplicar los criterios de elegibilidad previamente definidos. Las discrepancias fueron resueltas mediante discusión y consenso entre ambos revisores, garantizando transparencia y reduciendo el sesgo de selección [11].

3. Resultados

Tras eliminar duplicados, se identificaron un total de 132 artículos potencialmente relevantes que fueron evaluados según los criterios de elegibilidad previamente definidos. Se incluyeron estudios que evaluaban el reconocimiento de la palabra mediante metodologías orientadas al umbral de máxima discriminación (UMD), comparaban modalidades oral y digital, incluían información sobre calibración o automatización, ofrecían acceso al texto completo y aportaban datos de confiabilidad, validez o correlatos fisiológicos pertinentes para la práctica clínica.

De los 132 registros iniciales, 18 fueron evaluados a texto completo. Durante esta fase, se excluyeron 8 artículos que no comparaban modalidades oral versus digital, no describían procedimientos de calibración ni automatización, o no reportaban datos psicométricos de confiabilidad o validez. Finalmente, diez estudios cumplieron con los criterios metodológicos y conformaron la muestra final.

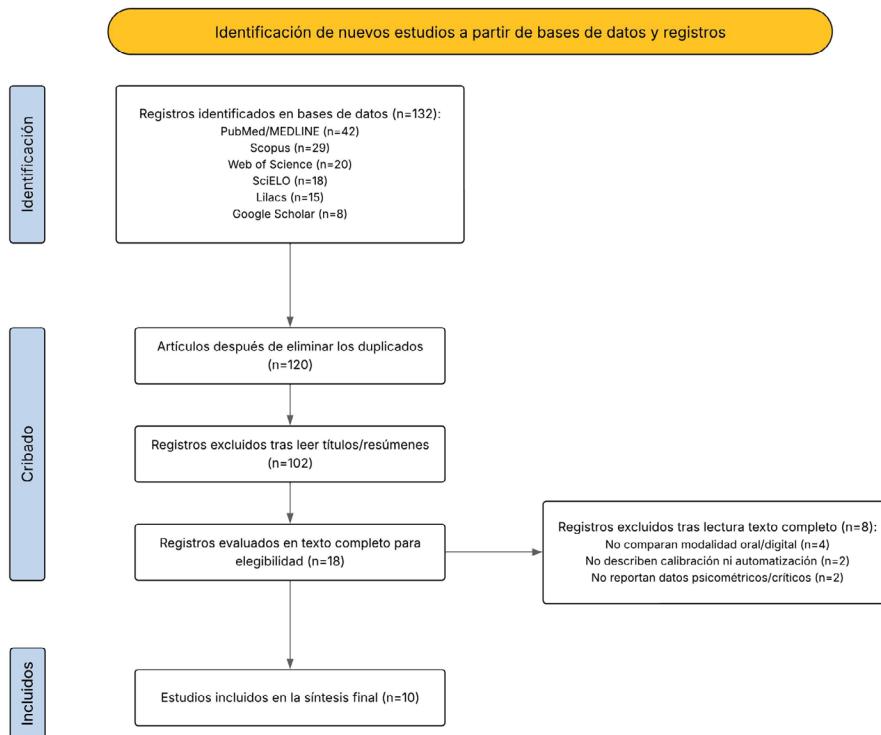
La Figura 1 presenta el diagrama de flujo PRISMA que resume las etapas del proceso: identificación, cribado, evaluación y selección.

Entre los estudios incluidos predominó el uso de listas digitalizadas: 8 emplearon exclusivamente modalidad digital, 1 comparó ambas modalidades (oral y digital), y solo 1 utilizó únicamente presentación oral tradicional. En consecuencia, el 90% de los estudios (9 de 10) utilizaron modalidad digital, y ninguno aplicó de manera sistemática el umbral PTP + 30 dB.

La Figura 1 presenta el diagrama de flujo tipo PRISMA que resume las etapas del proceso: identificación, cribado, evaluación y selección. Se documentó cada fase para asegurar la transparencia.

Figura 1.

Diagrama de flujo tipo PRISMA que representa el proceso de identificación, cribado, evaluación de elegibilidad e inclusión de estudios en la revisión de alcance.



La extracción de datos se realizó mediante una matriz estandarizada diseñada específicamente para capturar las variables de interés del estudio. Esta matriz incluyó información sobre características básicas del estudio (año, país, idioma, tipo de estudio, tamaño y características de la muestra), aspectos metodológicos específicos (modalidad de presentación, procedimientos de calibración, automatización), y resultados relacionados con análisis psicométricos, confiabilidad y validez (Tabla 1). La estandarización de este proceso permitió la comparación sistemática entre estudios y facilitó la identificación de patrones y brechas en la literatura.

El análisis de los datos extraídos adoptó un enfoque cualitativo-descriptivo, reconociendo que la heterogeneidad metodológica de los estudios incluidos no permitía la realización de meta-análisis cuantitativos [10]. En su lugar, se optó por una síntesis narrativa que permitiera identificar tendencias, convergencias y divergencias en la evidencia disponible, así como la caracterización de las principales brechas en el conocimiento actual.

Tabla 1.
Síntesis de las principales características metodológicas de los estudios incluidos en la revisión de alcance.

#	Referencia	Tipo de estudio	Idioma	Población/ muestra	Modalidad Oral viva	Modalidad Digital	Automatización / software	UMD fijo (PTP+30 dB)	Calibración detallada	Análisis fiabilidad
1	Ciscare et al., 2020 [4]	Experimental	Portugués	Niños (n=30)	Sí	Sí	Sí	No	Sí	Sí
2	Carlo et al., 2020 [3]	Prospectivo descriptivo	Español (Puerto Rico)	Adultos (n=50)	No	Sí	Sí	No	Sí	Sí
3	Sanchez et al., 2022 [13]	Revisión	Inglés	No aplica	No	Sí	Sí	No	No	Sí
4	Rodríguez-Ferreiro et al., 2024 [6]	Transversal descriptivo	Español	Adultos con hipoacusia (n=104)	No	Sí	Sí	No	No	Sí
5	Billings et al., 2023 [2]	Descriptivo, comparación entre tests	Inglés	Adultos y mayores (n=30)	No	Sí	Sí	No	Sí	Sí
6	Paz-Oliveira et al., 2020 [16]	Revisión integrativa	Portugués	No aplica	No	Sí	Sí	No	No	Sí
7	Rodríguez-Ferreiro & Serra, 2024 [12]	Revisión sistemática	Español	No aplica	No	Sí	Sí	No	No	Sí
8	Alokaily et al., 2022 [7]	Piloto (validación herramienta)	Inglés	Adultos normales (n=30)	No	Sí	Sí	No	Sí	Sí
9	Marcotti et al., 2021 [17]	Revisión conceptual	Español	No aplica	No	Sí	Sí	No	No	No
10	Páez Pimilla et al., 2024 [18]	Piloto descriptivo	Español	Adultos hipoacusia (n=60)	Sí	No	Sí	No	No	Sí

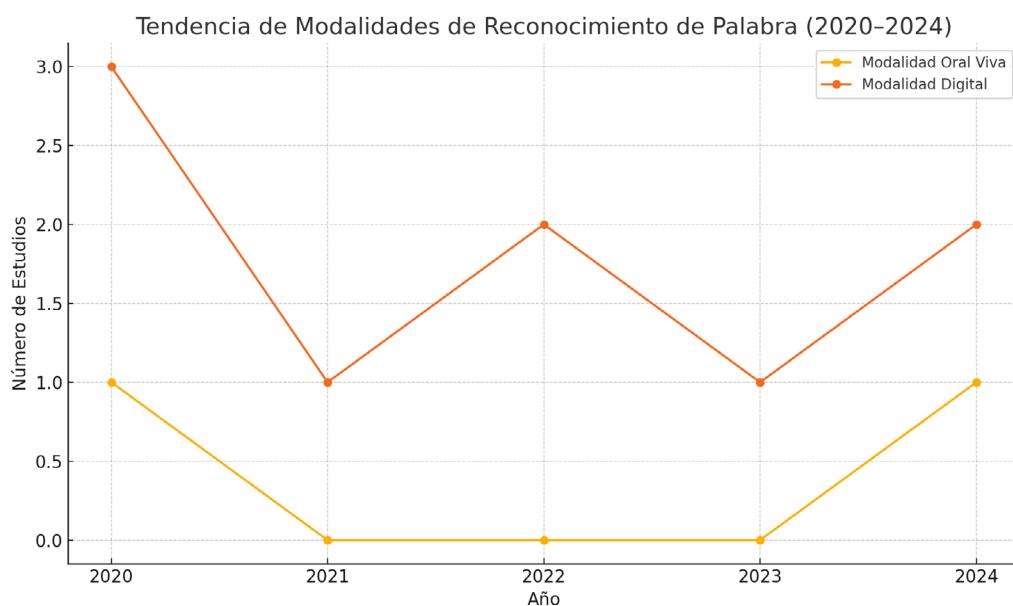
El análisis de las características metodológicas demostró una distribución heterogénea pero representativa de los diferentes enfoques investigativos en el campo. Los estudios incluidos abarcaron diversos diseños metodológicos, incluyendo investigaciones descriptivas, estudios piloto, revisiones sistemáticas y conceptuales, lo que refleja la naturaleza exploratoria y emergente del área de investigación. En términos de distribución lingüística, predominaron los artículos publicados en español e inglés, con menor representación en portugués, evidenciando el alcance multicultural de la temática pero también la concentración de la producción científica en determinadas regiones geográficas.

La caracterización de las poblaciones estudiadas mostró cierta diversidad, incluyendo adultos con hipoacusia, adultos con audición normal, adultos mayores y población pediátrica. Esta heterogeneidad poblacional, aunque positiva en términos de representatividad, también evidencia la necesidad de ampliar las investigaciones hacia grupos etarios y clínicos más diversos para consolidar la evidencia en diferentes contextos demográficos y patológicos.

La Figura 2 ilustra la evolución del uso de modalidades oral viva y digital en estudios de reconocimiento de la palabra entre 2020 y 2024 (años de los artículos incluídos en esta revisión). Se observa una predominancia de la modalidad digital, con presencia continua en todos los años analizados, alcanzando su punto más alto en el año 2020 y manteniéndose estable hasta 2024. En contraste, la modalidad oral muestra una tendencia decreciente, apareciendo únicamente en 2020-2024, y ausente durante tres años consecutivos.

Figura 2.

Diagrama de tendencias: evolución anual de publicaciones (2020–2024) sobre listas digitalizadas vs. orales en pruebas de reconocimiento de la palabra.



En conjunto, los resultados permiten observar una tendencia creciente hacia la adopción de listas de palabras en formato digital, evidenciada por el hecho de que nueve de los diez estudios incluidos emplearon algún tipo de modalidad automatizada. Esta frecuencia sugiere un interés progresivo por incorporar herramientas tecnológicas en la evaluación audiológica, particularmente aquellas orientadas a mejorar la estandarización y control de variables.

4. Discusión

Los resultados de esta revisión muestran que la digitalización de las listas de palabras está experimentando un crecimiento sostenido en la evaluación audiológica. Esta preferencia responde a la necesidad de resolver problemas históricos de la presentación oral como las variaciones en entonación, articulación o fatiga vocal del evaluador que afectan inevitablemente la consistencia de las mediciones. Al estandarizar acústicamente los estímulos, la digitalización permite alcanzar una mayor uniformidad entre diferentes profesionales y centros clínicos. No obstante, es importante ser prudentes, pues aunque los beneficios parecen prometedores, la evidencia disponible sigue siendo variable y aún insuficiente para establecer conclusiones firmes.

Diversos estudios han incorporado herramientas tecnológicas avanzadas, como aplicaciones móviles, plataformas web con aleatorización automática, control preciso de intensidad y almacenamiento en la nube. Estos avances reflejan un esfuerzo sostenido por modernizar la práctica audiológica; sin embargo, su implementación enfrenta barreras estructurales como el acceso desigual a dispositivos y conectividad, la falta de capacitación específica del personal y, en particular, la escasa validación en poblaciones vulnerables. Si bien podría plantearse que las listas digitalizadas presentan dificultades para personas con menor familiaridad tecnológica, como adultos mayores o individuos con deterioro cognitivo, este no es un problema inherente al formato digital. Los mismos desafíos pueden observarse con la presentación oral tradicional, donde influyen factores como la comprensión de instrucciones, la atención sostenida o la fatiga. En este sentido, la diferencia no radica en el método, sino en la mediación profesional: es la competencia del evaluador la que garantiza la accesibilidad, la adaptación del procedimiento y la interpretación adecuada de los resultados, independientemente de la modalidad empleada [12].

Uno de los hallazgos más relevantes fue la ausencia casi total del umbral óptimo simplificado. Aunque algunos estudios utilizaron niveles constantes, ninguno aplicó este criterio de manera sistemática. Esto representa una oportunidad perdida, especialmente considerando que organizaciones como la American Speech-Language-Hearing Association (ASHA) y la British Society of Audiology (BSA) han enfatizado la importancia de estandarizar los niveles de presentación para mejorar la comparabilidad entre evaluaciones. Un protocolo homogéneo no sólo facilitaría la comparación entre estudios, sino que fortalecería la base empírica de la disciplina y permitiría establecer valores normativos más confiables.

Las descripciones metodológicas sobre calibración acústica fueron, en muchos casos, vagas o insuficientes, comprometiendo seriamente la replicabilidad. A esto se suma el uso predominante de análisis estadísticos básicos como pruebas t de Student o Wilcoxon, sin recurrir a herramientas más robustas como el coeficiente de correlación intraclass (ICC) o los gráficos de Bland-Altman, recomendados en las guías COSMIN para evaluar la confiabilidad de instrumentos clínicos. Estas técnicas habrían permitido determinar no solo si las diferencias entre formatos son estadísticamente significativas, sino si realmente tienen relevancia clínica [13,14].

Aunque este trabajo no se centró en evaluación neurofisiológica, ninguno de los estudios revisados integró mediciones objetivas complementarias como los potenciales evocados auditivos del tronco encefálico (PEATC) o las respuestas auditivas de estado estable (ASSR). La incorporación de estas herramientas en investigaciones futuras aportaría evidencia objetiva que refuerce los hallazgos conductuales y contribuya a una comprensión más profunda del procesamiento auditivo [2].

Desde el punto de vista lingüístico, persiste la necesidad de validación transcultural rigurosa. Aunque hay esfuerzos por adaptar listas a distintos idiomas, las diferencias prosódicas y fonéticas entre

lenguas exigen procesos sistemáticos que aseguren la equivalencia de los estímulos. La International Test Commission (ITC) ha establecido directrices claras para la adaptación de pruebas entre culturas, pero su aplicación en el ámbito audiológico sigue siendo escasa [15, 16].

En conjunto, esta revisión evidenció una marcada heterogeneidad metodológica entre los estudios, tanto en los criterios de selección de participantes como en los procedimientos de calibración y análisis estadístico. Esta diversidad impidió realizar un meta-análisis cuantitativo con resultados comparables. Además, no puede descartarse la existencia de sesgos de idioma o de publicación, dado que la mayoría de los artículos revisados fueron publicados en inglés y en revistas con orientación clínica occidental, lo cual podría limitar la representatividad de los hallazgos a nivel global.

En consonancia con esto, investigaciones recientes han propuesto la digitalización de estímulos grabados oralmente y procesados, con el objetivo de controlar las condiciones acústicas y reducir la variabilidad asociada a la presentación oral en vivo [17,18]. Estos enfoques comparten con las listas digitalizadas de reconocimiento de palabras la meta de aumentar la confiabilidad diagnóstica y la comparabilidad entre evaluaciones. La discusión de estos enfoques en los resultados de esta revisión refuerza la pertinencia de avanzar hacia protocolos unificados que integren los principios metodológicos comunes de estandarización y control acústico, independientemente de las diferencias instrumentales o de formato.

En conjunto, esta revisión evidenció una marcada heterogeneidad metodológica entre los estudios, tanto en los criterios de selección de participantes como en los procedimientos de calibración y análisis estadístico. Esta diversidad impidió realizar un metaanálisis cuantitativo con resultados comparables. Además, no puede descartarse la existencia de sesgos de idioma o de publicación, dado que la mayoría de los artículos revisados fueron publicados en inglés y en revistas con orientación clínica occidental, lo cual podría limitar la representatividad de los hallazgos a nivel global.

5. Implicaciones para la práctica clínica y la investigación futura

Los hallazgos de esta revisión trazan un camino para avanzar en este campo. Es urgente validar el uso del umbral óptimo simplificado como estándar en la presentación digitalizada de listas de palabras. Esto requiere estudios multicéntricos que evalúen su aplicabilidad en diferentes poblaciones y contextos clínicos.

Los estudios futuros deberán fortalecer su rigor metodológico mediante descripciones exhaustivas de la calibración acústica, uso de análisis estadísticos adecuados para evaluar concordancia clínica (como ICC y Bland-Altman) y la integración de mediciones neurofisiológicas complementarias. Estas mejoras son necesarias para que la evidencia generada sea verdaderamente útil en la toma de decisiones clínicas.

El desarrollo de herramientas digitales en audiología debe equilibrar la precisión técnica con la viabilidad real de su aplicación. No basta con innovar, es necesario crear soluciones accesibles, sostenibles y adaptadas a las condiciones de cada sistema de salud, especialmente en contextos donde los recursos son limitados y la brecha digital aún marca diferencias profundas. La digitalización ofrece un potencial transformador para la evaluación audiológica, pero ese potencial solo se concretará si se sustenta en evidencia científica sólida, reproducible y pertinente a la diversidad de escenarios clínicos del mundo real.

6. Conclusiones y recomendaciones

Los resultados de esta revisión confirman que la digitalización de las listas de palabras constituye una alternativa prometedora para mejorar la estandarización y reproducibilidad en la evaluación del reconocimiento de la palabra. Sin embargo, la falta de consenso en los niveles de presentación y la variabilidad metodológica entre estudios impiden, por ahora, establecer un protocolo universal.

Avanzar hacia la estandarización del umbral óptimo simplificado y de la digitalización de palabras fortalecerá la comparabilidad internacional de las pruebas y marcará un paso decisivo hacia una audiología más precisa, equitativa y universal.

Referencias bibliográficas

1. Austin A, Ladner K, Mendel LL. Recorded word recognition testing is worth the time. *Am J Audiol.* 2025; 34(1): 179-186. https://doi.org/10.1044/2024_AJA-24-00080
2. Billings CJ, Olsen TM, Charney L, Madsen BM, et al. Speech-in-noise testing: An introduction for audiologists. *Semin Hear.* 2023; 45(1): 55–82. <https://doi.org/10.1055/s-0043-1770155>
3. 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.* 2020; 31(7): 531–546. <https://doi.org/10.1055/s-0040-1709446>
4. Ciscare GKSS, Zabeu JS, dos Santos DR, Morettin-Zupelari M, et al. List of words to evaluate speech perception: Recording and verification of applicability. *Revista CEFAC.* 2020; 22(5): e2820. <https://doi.org/10.1590/1982-0216/20202252820>
5. Margolis RH, Wilson RH, Saly GL, Gregoire HM, et al. Automated Forced-Choice Tests of Speech Recognition. *J Am Acad Audiol.* 2021; 32(9): 606–615. <https://doi.org/10.1055/s-0041-1733964>
6. Rodríguez-Ferreiro M, Durán-Bouza M, Marrero-Aguiar V. Analysis of the Spanish Auditory Test of Speech in Noise (PAHRE) in a population with hearing loss. *Audiol Res.* 2024; 14(5): 861–874. <https://doi.org/10.3390/audiolres14050073>
7. Alokaily AO, Alqabbani AF, Aleid A, Alhussaini K. Toward accessible hearing care: The development of a versatile Arabic word-in-noise screening tool: A pilot study. *Appl Sci.* 2022; 12(23): 12459. <https://doi.org/10.3390/app122312459>
8. Wilson RH, Scherer NJ. A quantitative protocol for calibrating short speech signals (monosyllabic words) based on the 50-ms segment of the voiced phoneme(s) with the maximum root-mean-square amplitude. *J Am Acad Audiol.* 2025; 36(2): 68–94. <https://doi.org/10.3766/jaaa.21126>
9. Arksey H, O’Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005; 8(1): 19–32. <https://doi.org/10.1080/1364557032000119616>
10. Tricco AC, Lillie E, Zarin W, O’Brien KK, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med.* 2018; 169(7): 467–473. <https://doi.org/10.7326/M18-0850>
11. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev.* 2016; 5(1): 210. <https://doi.org/10.1186/s13643-016-0384-4>

12. Rodríguez-Ferreiro M, Serra V. Pruebas de habla en ruido: una revisión de las pruebas disponibles en español. *Auditio*. 2024; 8: e113. <https://doi.org/10.51445/sja.auditio.vol8.2024.113>
13. Sanchez VA, Arnold ML, Moore DR, Clavier O, et al. Speech-in-noise tests: Innovative applications for pediatric patients, underrepresented populations, job fitness, clinical trials, and remote services. *J Acoust Soc Am*. 2022; 152(4): 2336. <https://doi.org/10.1121/10.0014418>
14. Zapala D, Stamper GC, Bogle JM, Jagger SL, et al. Clinical utility of the standardized word recognition score. *Ear Hear*. 2024; 45(1): 94–105. <https://doi.org/10.1097/AUD.0000000000001404>
15. Vainutienė V, Ivaška J, Kardelis V, Ivaškienė T, et al. Speech Audiometry: The Development of Lithuanian Bisyllabic Phonemically Balanced Word Lists for Evaluation of Speech Recognition. *Appl Sci*. 2024; 14(7): 2897. <https://doi.org/10.3390/app14072897>
16. Paz-Oliveira A, Momensohn-Santos TM, do Carmo MP, Fiore A. Testes de fala no ruído na clínica audiológica – Uma revisão integrativa. *Distúrb Comun*. 2020; 32(1): 124–139. <https://doi.org/10.23925/2176-2724.2020v32i1p124-139>
17. Marcotti FA, Galaz M, Iturriaga O, Aguilar S. Pruebas monoaurales de habla de baja redundancia: Evaluación de la separación/cierre monoaural. *Rev. Otorrinolaringol. Cir. Cabeza Cuello*. 2021; 81(2): 306-318. <https://doi.org/10.4067/s0718-48162021000200306>
18. Páez Pinilla ATP, Arjona CM, Montiel KPM, Avendaño Díaz EMA. Prueba de discriminación de habla en ruido: Etapa 2: Pilotaje en sujetos con pérdida auditiva neurosensorial simétrica leve a moderada. *Areté*. 2024; 24(2): 41-49. <https://doi.org/10.33881/1657-2513.art.24205>

Conflictos de intereses

Los autores no declaran conflictos de intereses con la realización de esta investigación.

Contribución de autoría

Génesis Olguín Hurtado: Conceptualización, análisis formal, investigación, metodología, Supervisión, Redacción – borrador original, Redacción – revisión y edición.

Mariano Blake: Conceptualización, análisis formal, investigación, metodología, Supervisión, Redacción – borrador original, Redacción – revisión y edición.

ABSTRACT

The evaluation of word recognition plays a key role in contemporary audiology, amid a growing debate on the advantages of digitized testing compared with traditional live-voice presentation. This scoping review analyzed studies published between 2020 and 2025, following the PRISMA-ScR methodology, to compare both modalities and to examine the use of the simplified optimal presentation level (PTA + 30 dB).

Findings indicate that digitized word lists provide greater reliability, repeatability, and inter-examiner consistency than live-voice presentation. Significant technological progress has been observed, including the development of mobile applications that promote automation and standardization of evaluations. However, major limitations persist, such as the limited adoption of the fixed-intensity criterion (PTA + 30 dB) and the absence of robust statistical analyses, including Bland-Altman plots and intraclass correlation coefficients.

Although digitalization has solid methodological foundations, further research employing rigorous comparative designs, advanced reliability analyses, and validation through neurophysiological correlates is required to consolidate its implementation as a clinical standard.

Keywords: Mobile Applications; Systematic Reviews as Topic; Audiology; Speech Recognition Software; Word Association Tests; Semantics.
