

Inteligencia artificial en audífonos: evidencia, aplicaciones clínicas e integración en el sector.

Victoria Eloísa Rodríguez Unamuno

SAERA. School of Advanced Education Research and Accreditation

RESUMEN

La hipoacusia afecta a más de 1.500 millones de personas en el mundo (Organización Mundial de la Salud [OMS], 2021), y los audífonos constituyen la intervención rehabilitadora de referencia. La incorporación de la inteligencia artificial (IA), en particular el aprendizaje automático (AM) y el aprendizaje profundo (AP), representa un cambio cualitativo en las posibilidades de procesamiento adaptativo, personalización y seguimiento del usuario. El objetivo de este trabajo fue analizar y sintetizar críticamente la literatura de acceso abierto sobre aplicaciones de IA en audífonos. Se llevó a cabo una revisión bibliográfica narrativa estructurada, en lugar de sistemática, dada la heterogeneidad metodológica de la evidencia disponible, aplicando criterios explícitos de trazabilidad. Los resultados se organizaron en cuatro ejes: reducción de ruido mediante AP con beneficio inicial en usuarios reales pero condicionado por las restricciones del dispositivo; reconocimiento de entornos acústicos para automatización con brecha entre precisión técnica y resultado clínico; personalización mediante comparaciones A/B y optimización bayesiana; y evaluación momentánea ecológica (EMA) como herramienta para capturar la experiencia auditiva en vida diaria. Se discuten implicaciones para la adaptación, teleaudiología y consideraciones éticas sobre privacidad, sesgos y explicabilidad. La revisión evidencia una brecha persistente entre la documentación de fabricantes y la validación científica independiente, y concluye que la IA puede aportar mejoras clínicamente relevantes, pero su impacto depende de la evaluación con usuarios reales y de la comunicación transparente de sus limitaciones.

Palabras clave: *inteligencia artificial, audífonos, aprendizaje automático, aprendizaje profundo, rehabilitación auditiva, evaluación momentánea ecológica, personalización.*

ABSTRACT

Hearing loss affects more than 1.5 billion people worldwide (World Health Organization [WHO], 2021), and hearing aids are the primary rehabilitative intervention. The incorporation of artificial intelligence (AI) —particularly machine learning (ML) and deep learning (DL)— represents a qualitative shift in the possibilities for adaptive processing, personalization, and user monitoring. This study aimed to critically appraise and synthesize open-access literature on AI applications in hearing aids. A structured narrative review, rather than a systematic one given the methodological heterogeneity of the available evidence, was

conducted with explicit traceability criteria, systematically separating peer-reviewed evidence from manufacturer technical documentation. Results were organized around four thematic axes: deep learning-based noise reduction with initial evidence of benefit in real users despite device constraints; acoustic scene recognition for automatic processing with a gap between technical accuracy and clinical outcomes; preference-based personalization through A/B comparisons and Bayesian optimization; and ecological momentary assessment (EMA) for capturing real-life auditory experience. Discussion addresses clinical implications for fitting and follow-up, teleaudiology integration, and ethical considerations regarding privacy, algorithmic bias, and explainability. A persistent gap between manufacturer technical documentation and independent scientific evidence was identified. The review concludes that AI can provide clinically meaningful improvements in hearing rehabilitation, but its impact depends on evaluation with real users and on transparent communication of its limitations.

Keywords: *artificial intelligence, hearing aids, machine learning, deep learning, hearing rehabilitation, ecological momentary assessment, personalization*

INTRODUCCIÓN

La audición constituye uno de los pilares fundamentales de la comunicación humana y la participación social. La Organización Mundial de la Salud (OMS) define la pérdida de audición como una capacidad auditiva inferior a la considerada normal, situando el umbral auditivo normal en ambos oídos en valores iguales o mejores que 20 dB; cuando el umbral supera este valor en una o ambas frecuencias del rango conversacional, se habla de hipoacusia, que puede clasificarse en leve, moderada, moderadamente grave, grave o profunda en función del grado de pérdida (OMS, 2024). Su deterioro, con independencia de la causa, tiene consecuencias que van más allá del plano perceptivo. La hipoacusia no tratada se asocia con mayor esfuerzo cognitivo durante la escucha, dificultades en la comprensión del habla, reducción de la participación social y un riesgo incrementado de deterioro cognitivo, identificado por la Lancet Commission (2024) como el principal factor de riesgo modificable para la demencia desde la mediana edad (Livingston et al., 2024). Según estimaciones de la OMS, más de 1.500 millones de personas presentan algún grado de pérdida auditiva en la actualidad, y las proyecciones apuntan a que esta cifra superará los 2.500 millones hacia 2050, situando la hipoacusia entre los problemas de salud de mayor prevalencia a escala global (OMS, 2021).

Frente a este panorama, los audífonos han sido históricamente la intervención rehabilitadora de referencia para la hipoacusia neurosensorial de leve a severa. Su evolución tecnológica ha sido notable.

Desde los primeros dispositivos analógicos hasta los sistemas digitales actuales, las ayudas auditivas han incorporado progresivamente compresión de rango dinámico, direccionalidad de micrófonos, reducción de ruido y conectividad inalámbrica, configurando plataformas de procesamiento en tiempo real orientadas a maximizar la audibilidad y el confort (Fabry & Bhowmik, 2021; Nada & Moaty, 2025). Sin embargo, y pese a estos avances, la comprensión del habla en condiciones de ruido y la satisfacción con la adaptación siguen siendo retos no resueltos para una proporción considerable de usuarios. Se ha estimado que aproximadamente la mitad de los usuarios de audífonos no se muestra satisfecho con el ajuste de sus dispositivos, si bien esta cifra debe interpretarse con cautela, ya que se refiere específicamente a la insatisfacción con la adaptación, y no necesariamente a una proporción equivalente de usuarios con fracaso en la comprensión del habla en ruido; aun así, las dificultades en entornos acústicos complejos constituyen uno de los factores que contribuyen a dicha insatisfacción (Balling et al., 2021; Nada & Moaty, 2025).

Los audífonos digitales contemporáneos son sistemas de procesamiento en tiempo real diseñados para convertir una audición reducida en audibilidad funcional y confortable. Para ello combinan prescripción de ganancia, compresión de rango dinámico, direccionalidad de los micrófonos y reducción de ruido, configurando una cadena que puede mejorar de forma apreciable el acceso al habla en condiciones controladas. Sin embargo, la escucha cotidiana rara vez se comporta como un escenario estable: en

presencia de ruido fluctuante, múltiples interlocutores, reverberación o cambios rápidos de contexto, la tarea auditiva deja de ser un problema de nivel y pasa a depender de procesos de discriminación de fuentes y atención selectiva. En consecuencia, la comprensión del habla en ruido continúa siendo una limitación reportada de forma consistente, incluso con adaptaciones técnicamente correctas.

A esta complejidad acústica se suma una complejidad humana significativa. Usuarios con perfiles audiométricos similares pueden diferir de manera marcada en tolerancia al ruido, demandas comunicativas, preferencias sonoras, expectativas y estilos de vida. Esta heterogeneidad hace que un ajuste “promedio” sea, con frecuencia, insuficiente y refuerza la necesidad de estrategias que permitan personalización y adaptación contextual sin incrementar el esfuerzo de uso (Jensen et al., 2019; Tasnim et al., 2024). En audiolología, además, el valor clínico de una tecnología no se agota en la mejora de métricas objetivas: la rehabilitación auditiva se evalúa por su impacto en la vida diaria, incluyendo participación, fatiga asociada al esfuerzo de escucha, satisfacción y continuidad de uso (Holube et al., 2020; Vercammen et al., 2023). Frente a esta doble complejidad, acústica e individual, la inteligencia artificial se plantea como una vía de respuesta, sobre la que ya existen síntesis orientadas al ámbito clínico hispanohablante (Cabrera et al., 2024).

Conceptos operativos: IA, AM y AP en audífonos

En este trabajo, la IA se emplea como etiqueta funcional para el conjunto de

técnicas computacionales que permiten inferir estados u optimizar decisiones de procesamiento a partir de datos. Dentro de ese marco, se distinguen el aprendizaje automático (AM), y el aprendizaje profundo (AP), con un propósito estrictamente operativo: facilitar el análisis comparativo de enfoques según cómo codifican su representación de entrada, su rol en la arquitectura del audífono y sus restricciones de implementación en un sistema integrado en tiempo real.

El AM se entiende aquí como el uso de modelos que se entrenan con descriptores acústicos definidos de forma explícita para abordar tareas de decisión, discriminación o clasificación. En audífonos, estos modelos suelen actuar como módulos de control que estiman probabilidades o etiquetas de contexto, p. ej., “habla en ruido”, “ruido estacionario”, “música”, “viento” y controlan parámetros o combinaciones de algoritmos ya presentes en el dispositivo (Hasemann & Krylova, 2024; Phonak, 2025).

El AP se utiliza para describir arquitecturas neuronales profundas que aprenden representaciones jerárquicas y realizan transformaciones a nivel de señal. En audífonos, ello incluye modelos de cálculo de máscaras o ganancias implementados mediante redes neuronales profundas (RNP), así como modelos de detección contextual. El AP se examina con atención especial a su compatibilidad con restricciones del dispositivo: procesamiento en tiempo real, baja latencia, eficiencia computacional, consumo energético y estabilidad perceptiva (Andersen et al., 2021; Schröter et al., 2020).

La distinción entre AM y AP no se entiende como una oposición rígida, sino como un recurso para ordenar las soluciones según el papel que desempeñan en el sistema. Esta revisión adopta una lectura por tareas —qué problema resuelve— y por restricciones —en qué condiciones es viable— más que por etiquetas comerciales.

Finalmente, estas definiciones operativas se vinculan con criterios centrados en la persona: se consideran tecnológicamente pertinentes aquellas aproximaciones cuya evaluación incluye, directa o indirectamente, consecuencias sobre la satisfacción, la adherencia, el esfuerzo o fatiga auditiva y el desempeño comunicativo en situaciones reales (Balling et al., 2021; Tasnim et al., 2024).

OBJETIVOS

Objetivo general

Analizar, organizar y sintetizar críticamente la literatura de acceso abierto sobre aplicaciones de IA en audífonos, evaluando su relevancia para la audiolología clínica y su posible impacto en una rehabilitación auditiva orientada a las necesidades de la persona usuaria.

Objetivos específicos

1. Analizar enfoques de IA aplicados a la reducción de ruido y el realce del habla en audífonos, identificando los compromisos entre beneficio potencial y viabilidad de implementación —latencia, eficiencia computacional, consumo y estabilidad perceptiva— a partir de la evidencia disponible.

2. Revisar la evidencia sobre el reconocimiento de entornos acústicos y su aplicación para automatizar el procesamiento, y examinar los métodos de personalización basados en las preferencias del usuario, comparaciones A/B y optimización bayesiana, atendiendo tanto al rendimiento como a las implicaciones funcionales y de usabilidad.

3. Describir la traslación de estas tecnologías a productos comerciales a partir de documentación técnica de acceso público, diferenciando de forma explícita entre declaraciones de rendimiento, validación técnica interna y evidencia revisada por pares.

4. Discutir las implicaciones clínicas, organizativas y éticas para la rehabilitación auditiva, incluyendo la adaptación, el seguimiento, la educación del usuario, la teleaudiología, la evaluación ecológica en vida diaria y la transparencia en el uso de sistemas de decisión automatizada.

Delimitación y criterios operativos del trabajo

Este trabajo se centra en los audífonos y sistemas de asistencia auditiva equivalentes. Las tecnologías implantables como los implantes cocleares (IC) quedan fuera del foco principal del análisis, aunque se mencionan de forma puntual cuando su evidencia resulta metodológicamente transferible al ámbito de los audífonos.

En cuanto a las fuentes, se han considerado dos categorías diferenciadas: publicaciones científicas de acceso abierto con revisión por pares (peer-reviewed) y documentación técnica pública de

fabricantes: informes técnicos, fichas técnicas y folletos con contenido verificable. Esta distinción es funcional, ambas fuentes aportan información relevante, pero desde niveles de evidencia distintos que el análisis mantiene separados en todo momento.

El eje de evaluación prioriza resultados con significado clínico o de uso — preferencia, satisfacción, adherencia, medidas ecológicas y desempeño en condiciones realistas— sobre métricas puramente instrumentales cuando estas no se vinculan a un impacto funcional demostrado. Toda afirmación sobre “mejora” se considera dependiente del escenario acústico, del perfil de usuario y del criterio de medida empleado, evitando generalizaciones que excedan las condiciones en que se han obtenido los datos.

METODOLOGÍA

Diseño y enfoque

Se llevó a cabo una revisión bibliográfica narrativa estructurada con criterios explícitos de trazabilidad. Se escogió el formato narrativo porque la evidencia sobre IA en audífonos es metodológicamente heterogénea: conviven estudios de ingeniería con señales grabadas, preimpresiones con propuestas de arquitectura, evaluaciones perceptuales con oyentes y trabajos con resultados clínicos o ecológicos. El objetivo fue integrar y comparar hallazgos en función de la tarea y de las restricciones de implementación propias del dispositivo como latencia, causalidad, consumo, memoria y robustez.

La elección de un enfoque narrativo estructurado, frente a una revisión sistemática siguiendo el protocolo PRISMA (Page et al., 2021), responde a tres razones concretas. En primer lugar, el objeto de estudio agrupa evidencia metodológicamente muy heterogénea — publicaciones revisadas por pares, preimpresiones con propuesta de arquitectura, documentación técnica de fabricantes y evaluaciones perceptuales—, lo que dificulta su homologación cuantitativa exigida por una revisión sistemática. En segundo lugar, el objetivo del trabajo no es responder a una pregunta clínica única que admita síntesis cuantitativa, sino integrar y comparar líneas de desarrollo tecnológico y clínico en función de su pertinencia para la práctica audiológica. En tercer lugar, la rápida evolución del campo y la escasez de ensayos pragmáticos con usuarios reales aconsejan una aproximación interpretativa que pondere críticamente el nivel de evidencia de cada fuente, en lugar de aplicar criterios de exclusión estrictos que descartarían documentación relevante. Por estos motivos, no se aplicó el protocolo PRISMA de forma formal. En su lugar, se establecieron criterios explícitos de elegibilidad, clasificación y trazabilidad que permiten identificar el origen y nivel de evidencia de cada afirmación incluida en la síntesis. El cribado cuantitativo no se planificó de forma sistemática durante el proceso; las cifras presentadas en el diagrama de flujo (Figura 1) corresponden a estimaciones a posteriori del recorrido aproximado de búsqueda y selección. La trazabilidad principal se garantiza mediante los criterios de elegibilidad explícitos (recogidos en la Tabla 5), la clasificación tipológica de cada fuente en

el Anexo A y la separación documental de los informes de fabricantes en el Anexo B.

Fuentes de información

Se incluyeron exclusivamente documentos con texto completo accesible. Como literatura científica de acceso abierto se consideraron artículos en repositorios biomédicos de texto completo (p. ej., PubMed Central), preimpresiones en repositorios abiertos (p. ej., arXiv) y comunicaciones de congresos con PDF libremente descargable. Como documentación técnica de acceso público de fabricantes se incluyeron informes técnicos, fichas técnicas y folletos descargables. Esta separación se mantuvo durante toda la síntesis para evitar equivalencias indebidas entre documentación corporativa y evidencia revisada por pares.

Las prepublicaciones y los documentos técnicos de fabricantes se utilizaron exclusivamente como fuentes de carácter exploratorio y descriptivo. Sus resultados no se equipararon a evidencia revisada por pares y se interpretan dentro de sus limitaciones metodológicas y de validación independiente. De forma puntual se incluyeron referencias fundacionales (p. ej., Cox & Alexander, 1995; Gatehouse & Noble, 2004; Stone & Moore, 2005; Stone et al., 2008) cuyo acceso íntegro requiere suscripción institucional pero cuyo contenido metodológico es indispensable para sostener afirmaciones sobre umbrales perceptivos y evaluación funcional.

Estrategia de búsqueda y recuperación.

La identificación de documentos se realizó mediante búsquedas por términos y

combinaciones orientadas a los cuatro ejes del trabajo, empleando palabras clave en inglés y español. La Tabla 2 recoge los términos empleados, organizados por eje temático.

De forma complementaria, se aplicaron sistemas de búsqueda inversa mediante la revisión de referencias citadas en artículos o documentos técnicos para comprender el método o el marco de evaluación y búsqueda por familias tecnológicas consistente en rastreo de documentos relacionados en repositorios públicos de fabricantes cuando un concepto específico era introducido por un producto concreto.

La relación completa de términos empleados se recoge en la Tabla 1, al final de esta sección.

Repositorios consultados y alcance

La búsqueda se distribuyó en cuatro tipos de fuentes complementarias, recogidas en la Tabla 3, con el fin de combinar evidencia científica revisada por pares y documentación de traslación a producto. Se consideraron documentos publicados hasta la fecha de cierre de la búsqueda (enero de 2026). El proceso detallado de búsqueda y selección se representa en la figura 1.

El detalle de los repositorios consultados figura en la Tabla 2, al final de esta sección.

Ejemplo representativo de ecuación de búsqueda

Para favorecer la reproducibilidad, las combinaciones de términos se formularon mediante operadores booleanos, agrupando los términos en tres bloques

conceptuales combinados entre sí. La Tabla 4 resume la estructura de la ecuación de búsqueda empleada.

La estructura completa de la ecuación de búsqueda se recoge en la Tabla 3, al final de esta sección.

Criterios de elegibilidad

Los criterios de inclusión y exclusión se presentan de forma detallada y comparativa en la Tabla 5. De manera resumida, se incluyeron documentos con aplicación directa a audífonos o a procesos vinculados a la rehabilitación auditiva, con uso explícito de IA, AM o AP y texto completo de acceso abierto. Para documentación de fabricantes, se requirió disponibilidad pública descargable y contenido técnico mínimo verificable.

Los criterios completos se presentan de forma comparativa en la Tabla 4, al final de esta sección.

Proceso de selección (trazabilidad)

La selección se realizó en tres etapas: (a) cribado inicial por título y resumen para confirmar relevancia; (b) verificación de elegibilidad por texto completo; y (c) clasificación por tipo de evidencia y asignación a ejes temáticos. Cada fuente quedó etiquetada como artículo científico (revisado por pares o preimpresión con método explícito) o documento técnico (informe técnico, ficha técnica o folleto de fabricante). La relación completa de documentos técnicos de fabricantes consultados se recoge en el Anexo B.

El recorrido completo del proceso de búsqueda y selección, desde la identificación inicial hasta la clasificación

final de documentos, se resume en la Figura 1. Aunque no se siguió el protocolo PRISMA de forma formal, por las razones metodológicas expuestas en el apartado de Diseño y enfoque, el diagrama respeta su lógica de trazabilidad.

El recorrido completo del proceso se representa gráficamente en la Figura 1, al final de esta sección.

Extracción de datos y esquema de codificación

Para homogeneizar la síntesis se realizó una extracción estructurada por documento, registrando: identificación; tarea principal; tipo de IA (AM, AP o híbrido); punto de integración; restricciones de implementación (causalidad, latencia, cómputo, consumo, robustez); tipo y nivel de evaluación; condiciones de validez; y naturaleza de las conclusiones. La Tabla 6 describe la plantilla de codificación empleada. Un ejemplo de ficha cumplimentada para el artículo de Schulz et al. (2026) se incluye en el Anexo C.

La plantilla de extracción empleada se detalla en la Tabla 5, al final de esta sección.

Consideraciones de calidad y manejo de limitaciones

La síntesis se guió por tres reglas: (a) adecuación de la inferencia al diseño: las mejoras en métricas instrumentales se interpretaron como evidencia de rendimiento técnico sin asumir beneficio funcional si no existía evaluación con usuarios; (b) separación de niveles de validación: los informes técnicos de fabricantes se utilizaron para describir

decisiones de implementación, pero sus afirmaciones de rendimiento se reportaron como evidencia técnica no revisada de forma independiente; y (c) dependencia del contexto: toda “mejora” se consideró condicionada al escenario acústico, al perfil de usuario y al criterio de evaluación. Como limitaciones metodológicas, se reconoce que restringir la revisión a texto completo abierto puede introducir sesgo de disponibilidad, y que el enfoque narrativo no persigue estimaciones agregadas comparables a una revisión sistemática.

Tablas y figuras de la sección de Metodología

Tabla 1.

Palabras clave empleadas en la búsqueda bibliográfica, organizadas por eje temático.

Eje temático	Términos en inglés	Equivalentes en español
Reducción de ruido y realce del habla	hearing aid(s), deep learning, machine learning, noise reduction, speech enhancement	audífonos, aprendizaje profundo, aprendizaje automático, reducción de ruido, realce del habla
Reconocimiento de escena y automatización	acoustic scene classification, environment recognition	clasificación de escena acústica, reconocimiento de entorno
Personalización por preferencias	preference learning, A/B comparison, Bayesian optimization	aprendizaje de preferencias, comparación A/B, optimización bayesiana
Teleaudiología y evaluación ecológica	remote fitting, teleaudiology, ecological momentary assessment	ajuste remoto, teleaudiología, evaluación momentánea ecológica

Nota. Los términos se combinaron mediante operadores booleanos (AND/OR) y se aplicaron tanto a búsquedas simples como a estrategias complementarias hacia atrás y por familias tecnológicas.

Tabla 2.

Repositorios y tipos de fuentes consultados durante la búsqueda bibliográfica.

Tipo de fuente	Ejemplos de repositorios	Función en la revisión
Repositorios biomédicos de acceso abierto	PubMed Central	Localización de artículos revisados por pares con texto completo
Repositorios de preimpresiones	arXiv	Localización de propuestas técnicas con método explícito
Actas y congresos con PDF descargable	IEEE (ICASSP, IWAENC)	Localización de contribuciones de ingeniería de señal
Repositorios y páginas oficiales de fabricantes	Phonak, Oticon, GN ReSound, Signia, Starkey, Widex	Documentación técnica de traslación industrial

Nota. Las búsquedas en repositorios biomédicos y de preimpresiones se realizaron mediante ecuaciones booleanas; las de repositorios de fabricantes se realizaron de forma dirigida cuando un concepto técnico era introducido por una plataforma comercial concreta.

Tabla 3.

Estructura de la ecuación de búsqueda booleana empleada.

Bloque	Operador interno	Términos
Dispositivo	OR	"hearing aid" OR "hearing aids"
Técnica de IA	OR	"deep learning" OR "machine learning" OR DNN
Aplicación	OR	"noise reduction" OR "speech enhancement" OR "acoustic scene" OR "environment recognition" OR "preference learning" OR "remote fitting" OR teleaudiology OR "ecological momentary assessment"

Nota. Los tres bloques se combinaron entre sí mediante el operador AND, de modo que cada documento recuperado contuviera al menos un término de cada bloque. La ecuación completa resultante fue: (Dispositivo) AND (Técnica de IA) AND (Aplicación).

Tabla 4.

Criterios de inclusión y exclusión aplicados en la selección de fuentes.

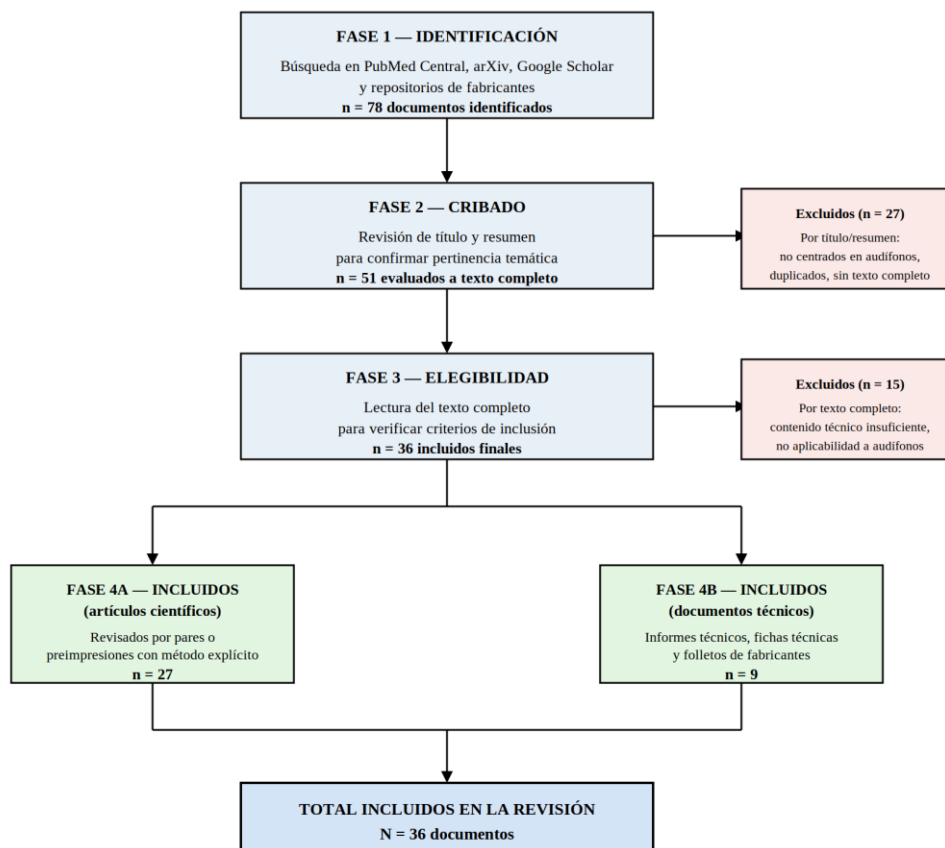
Criterios de inclusión	Criterios de exclusión
Aplicación directa a audífonos o a procesos estrechamente vinculados a la rehabilitación auditiva: ajuste, seguimiento, personalización, evaluación funcional o teleaudiología.	Ausencia de acceso al texto completo en cualquiera de las categorías de fuente consideradas.
Uso explícito de IA, AM o AP, con descripción técnica suficiente para identificar: la tarea que resuelve el modelo, la representación de entrada/salida y el punto de integración en el sistema.	Documentos no centrados en audífonos o sin transferibilidad razonable al contexto audiológico (p. ej., audio genérico sin discusión de restricciones embebidas).
Texto completo de acceso abierto: artículos en repositorios biomédicos (PubMed Central), preimpresiones (arXiv) o comunicaciones de congresos con PDF libremente descargable.	Duplicados: cuando la misma contribución aparece en múltiples versiones, se conservó únicamente la versión más completa y técnicamente informativa.

Criterios de inclusión	Criterios de exclusión
Excepción: referencias fundacionales (Cox & Alexander, 1995; Gatehouse & Noble, 2004; Stone & Moore, 2005; Stone et al., 2008) cuyo contenido metodológico es indispensable.	
Para documentación técnica de fabricantes: disponibilidad pública descargable y contenido técnico mínimo verificable (descripción de arquitectura/función, supuestos de diseño y, si existe, evaluación interna reportada).	Material exclusivamente promocional sin contenido técnico verificable. Excepción: informes técnicos de fabricante identificados como tales y clasificados explícitamente como documentación no revisada por pares.
Documentos publicados hasta la fecha de cierre de la búsqueda (enero de 2026), sin restricción de idioma cuando el texto completo era accesible.	Documentos cuyo foco principal son tecnologías implantables (implantes cocleares) sin discusión explícita de transferibilidad metodológica a audífonos convencionales.

Nota. IA = inteligencia artificial; AM = aprendizaje automático; AP = aprendizaje profundo.

Figura 1.

Diagrama de flujo del proceso de búsqueda y selección de fuentes.



Nota. Las cifras corresponden a estimaciones a posteriori del recorrido aproximado de búsqueda y selección, dado que no se realizó cribado cuantitativo formal por tratarse de una revisión narrativa estructurada (véase el apartado de Diseño y enfoque). El diagrama distingue documentos científicos revisados por pares de documentos técnicos de fabricantes, que reciben un tratamiento diferenciado en toda la síntesis.

Tabla 5.

Plantilla de extracción y codificación de fuentes (resumen operativo).

Campo	Qué se registra	Valores / guía de codificación
ID interno	Código único	Ej.: A01, A02... (artículos) / T01, T02... (técnicos)
Tipo de fuente	Naturaleza del documento	Artículo científico (revisado por pares), preimpresión, actas, documento técnico (informe técnico, ficha técnica o folleto de fabricante)
Dominio	Ámbito principal	Procesamiento de señal, AM/AP, clínica/rehabilitación, teleaudiología, evaluación ecológica
Tarea primaria	Problema que resuelve	(E1) Reducción de ruido/realce del habla, (E2) clasificación de escena/entorno, (E3) personalización por preferencias, (E4) soporte remoto/sensores/monitorización
Tarea secundaria	Si aplica	Misma lista (E1–E4)
Enfoque IA	Familia metodológica	AM, AP, híbrido
Representación de entrada	Qué consume el modelo	Características (basadas en descriptores), espectrograma/TF, señal temporal, metadatos/uso, multimodal (audio + sensores)
Salida del modelo	Qué produce	Etiqueta/probabilidad de escena, máscara/ganancia, señal realzada, parámetros recomendados, ranking de preferencias, política de ajuste
Punto de integración	Dónde actúa	Control/selección de programa, modulación de RR/direccionalidad/compresión, procesado de señal (etapa previa), aplicación/UX (retroalimentación A/B), servicio remoto/teleajuste

Campo	Qué se registra	Valores / guía de codificación
Causalidad	Acceso temporal	Causal / no causal / no especificado
Latencia	Información reportada	Latencia explícita (ms), latencia implícita (ventana), o no reportada
Restricción computacional	Viabilidad integrada	Modelo liviano, cuantización, requisitos hardware, consumo; si no hay datos: “no reportado”
Robustez	Generalización	Evaluación en condiciones diversas (sí/no); si no se reporta: “no evaluado”
Comparador	Con qué se compara	Referencia clásica (RR tradicional), sistema comercial, otro modelo IA, condición sin procesamiento, “sin comparador”
Tipo de evaluación	Nivel de evidencia	Objetiva (métricas), perceptual (oyentes), clínica (usuarios), ecológica (vida diaria/EMA), mixta
Métricas / resultados primarios	Qué se mide	Inteligibilidad, SNR/SDR/STOI/PESQ, preferencia, SSQ, satisfacción, adherencia, fatiga/esfuerzo, etc.
Población	Quién participa	Audición normal, hipoacusia, usuarios de audífono; si no aplica: “N/A”
Condiciones	Contexto experimental	Tipo de ruido, reverberación, escena, movilidad, niveles
Conclusión principal	Qué afirma el estudio	Frase breve, restringida al diseño y condiciones
Limitaciones declaradas	Lo que el propio texto reconoce	Generalización, tamaño muestral, condiciones, sesgos, etc.
Riesgo de sobreinterpretación	Nota del revisor	Salto de métricas objetivas a beneficio clínico sin soporte; o informe técnico sin validación independiente

RESULTADOS Y REVISIÓN TEMÁTICA

Los documentos identificados y clasificados conforme a los criterios establecidos se distribuyen en dos categorías: artículos científicos revisados por pares o preimpresiones con estructura científica, e informes técnicos de fabricantes con contenido técnico verificable. Los hallazgos se presentan en torno a los cuatro ejes temáticos del trabajo, con un eje transversal dedicado a la traslación industrial y a la ampliación del rol rehabilitador del audífono mediante sensores y funciones de salud. Cada eje temático se cierra con una síntesis estructurada en tres componentes: qué se sabe, qué falta e implicación clínica, con el fin de mantener explícita la conexión entre hallazgos técnicos y práctica audiológica. La selección de estudios clave incluidos en la revisión se presenta de forma tabulada en el Anexo A.

La evidencia disponible presenta una distribución desigual en términos de nivel metodológico. La mayoría de los estudios identificados corresponden a evaluaciones técnicas en condiciones de laboratorio o simulación, mientras que los estudios con usuarios reales en condiciones ecológicas son menos frecuentes (Nada & Moaty, 2025; Tasnim et al., 2024). Esta asimetría limita la extrapolación directa de los resultados técnicos al beneficio clínico en la vida diaria y refuerza la pertinencia de mantener separados, a lo largo de la síntesis, los niveles de validación de cada hallazgo.

Reducción de ruido y realce del habla

La reducción de ruido basada en modelos de AP representa uno de los avances más documentados en la aplicación de IA a los audífonos. A diferencia de los algoritmos clásicos fundamentados en estimación estadística del ruido, como la sustracción espectral o el filtrado de Wiener, los modelos de AP aprenden representaciones complejas de la señal que les permiten discriminar habla de ruido incluso en condiciones de superposición espectral elevada y variabilidad acústica de las escenas reales (Andersen et al., 2021; Aubreville et al., 2018). En enfoques tradicionales, el beneficio suele emerger de la combinación de direccionalidad y postfiltrado. Los modelos neuronales se han propuesto como complemento o sustituto parcial de los estimadores clásicos, con la intención de manejar mejor escenarios no estacionarios y escenas acústicas complejas. Esto solo resulta viable, sin embargo, si el procesamiento puede ejecutarse en tiempo real.

Las primeras propuestas en esta línea se orientaron a soluciones causales y de baja latencia. Aubreville et al. (2018) propusieron una red que predice una ganancia inspirada en el filtro de Wiener con un contexto temporal asimétrico, precisamente para limitar la latencia sin renunciar a información contextual útil. Schröter et al. (2020) introdujeron CLCNet, una arquitectura de codificación lineal compleja orientada a representar la señal compleja preservando la información de fase relevante para la naturalidad del sonido realzado. Andersen et al. (2021), en una revisión técnica integradora del estado del arte, describieron cómo los sistemas basados en

RNP pueden generar claridad en entornos ruidosos, ilustrando con ejemplos los compromisos entre métricas objetivas, latencia y valoraciones perceptuales en oyentes con hipoacusia.

Más recientemente, el enfoque se ha desplazado hacia evaluaciones con usuarios reales y arquitecturas distribuidas. Diehl et al. (2023) presentaron un sistema de reducción de ruido basado en AP cuyo procesamiento computacional se delega a un teléfono inteligente conectado inalámbricamente al audífono, describiendo no solo mejoras en habla en ruido en usuarios con hipoacusia moderada a severa (mejora de 1,6 dB en SRT y aumento de más del 40 % en valoraciones subjetivas frente a un audífono adaptado), sino también la latencia total del sistema (aproximadamente 65–75 ms según el modelo de teléfono utilizado), parámetro crítico para la interacción en tiempo real, que permite discutir no solo si el sistema mejora, sino si el resultado es compatible con la interacción cotidiana. Ashkanichenarlogh et al. (2025) aportaron evidencia en condiciones de ajuste y escucha diversas mediante maniquí acústico (KEMAR) y estimadores de inteligibilidad auditivamente informados, como el índice HASPI. Saoji et al. (2024) plantearon cómo los esquemas de reducción de ruido basados en RNP podrían influir en medidas relacionadas con la candidatura a implante coclear: en una muestra de 10 pacientes con hipoacusia severa (PTA promedio del oído mejor de 67 dB), la reducción de ruido basada en RNP mejoró la percepción del habla entre 20 y 32 puntos porcentuales respecto al programa basal, situando el

debate en un marco de toma de decisiones más amplio que la inteligibilidad aislada.

El trabajo más reciente del campo profundiza en este último aspecto extendiendo la evaluación a perfiles clínicamente relevantes. Schulz et al. (2026) evaluaron el efecto de un sistema de AP de latencia ultrabaja (1 ms, basado en Conv-TasNet) en tres grupos: 20 oyentes con audición normal (NH, del inglés normal hearing), 19 personas con hipoacusia (HI, del inglés hearing impairment) y 12 usuarios de implante coclear (IC) inscritos, de los cuales 10 fueron analizados por problemas de conexión con el streamer en 2 participantes, mostrando una mejora significativa en la mediana del umbral de recepción del habla (SRT) en HI (0,8 dB) y, marcadamente en IC (5,7 dB), junto con un ligero deterioro en NH (−0,9 dB). Vaisberg et al. (2025) incorporaron correlatos neurales del procesamiento auditivo en 26 usuarios experimentados de audífono, evaluando cambios en la oxigenación de la corteza prefrontal izquierda mediante espectroscopia de infrarrojo cercano funcional (fNIRS) durante tareas de escucha en ruido.

Qué se sabe. Los enfoques basados en AP pueden aportar ventajas en escenarios de ruido no estacionario y escenas complejas. La literatura refleja una transición desde “mejoras de separación” hacia diseños explícitamente orientados a viabilidad en tiempo real y a condiciones representativas del dispositivo.

Qué falta. Evidencia más consistente con usuarios reales y en condiciones ecológicas, con resultados que vayan más allá de la inteligibilidad: preferencia sostenida, fatiga o esfuerzo, aceptación a

largo plazo. También falta homogeneidad en comparadores y en el reporte de parámetros críticos como la latencia total efectiva, señales de un avance técnico que va por delante de su comprobación con usuarios reales.

Implicación clínica. La incorporación de AP en reducción de ruido debe interpretarse como una intervención rehabilitadora potencial, cuya utilidad depende tanto del beneficio auditivo como de su aceptabilidad diaria. En adaptación clínica, esto sugiere priorizar protocolos que incluyan verificación funcional en ruido, educación sobre expectativas y seguimiento de experiencia.

Restricciones específicas del dispositivo

La integración de modelos de AP en audífonos impone restricciones de implementación que condicionan de forma determinante tanto el diseño de las arquitecturas como la generalización de los resultados obtenidos en contextos de investigación. La restricción más crítica es la latencia. En términos generales, la latencia total del sistema debe mantenerse por debajo de aproximadamente 5–6 ms en adaptaciones auditivas abiertas (Stone et al., 2008) y por debajo de aproximadamente 20–30 ms en adaptaciones cerradas (Stone & Moore, 2005). Con ello se persigue evitar el efecto de peine, una interferencia destructiva entre la señal acústica directa y la señal procesada que se percibe como coloración o distorsión. Estos valores constituyen umbrales de tolerancia perceptiva ampliamente referidos en la literatura y son los que adoptan, como criterio de viabilidad, los desarrollos recientes en AP (Schulz et al., 2026). Muchas propuestas

adoptan diseños centrados en inferencia causal o casi causal, ventanas cortas y modelos compactos, priorizando estabilidad perceptiva y viabilidad embebida.

Diversos trabajos han propuesto estrategias específicas para conciliar rendimiento y restricciones embebidas. Schröter et al. (2020) abordaron esta restricción mediante la representación compleja de la señal, logrando procesamiento causal eficiente sin renunciar a la información de fase. Schulz et al. (2026) diseñaron su evaluación explícitamente en torno a modelos de baja latencia compatibles con el uso clínico, situando el algoritmo en 1 ms (16 muestras a 16 kHz). Además de la latencia, los audífonos operan bajo restricciones de consumo energético, memoria y capacidad de cómputo que limitan la complejidad de los modelos integrados. Las soluciones exploradas incluyen la cuantización de pesos, el podado de conexiones y el diseño de arquitecturas ligeras (Aubreville et al., 2018). La alternativa del procesamiento distribuido vía teléfono inteligente (Diehl et al., 2023) desplaza el problema computacional fuera del dispositivo, pero introduce dependencias de conectividad que pueden afectar la usabilidad y elevar la latencia total a 65–75 ms.

Qué se sabe. El “estado del arte” en AP para audífonos no se define solo por el rendimiento de separación, sino por cómo el modelo negocia latencia, estabilidad y robustez. Las restricciones embebidas condicionan de forma determinante la transferibilidad de los resultados de investigación al producto clínico.

Qué falta. Mayor detalle y homogeneidad en el reporte de parámetros de

implementación: latencia total efectiva del sistema, consumo energético, requisitos de memoria y comportamiento en condiciones de ruido no visto.

Implicación clínica. Los resultados de estudios sin validación en hardware real o sin consideración explícita de restricciones de dispositivo deben interpretarse con cautela. La formación del audiólogo en este aspecto facilita una lectura crítica de la evidencia, sin dar por hecho que el rendimiento en laboratorio se traslade igual al día a día.

Evidencia en usuarios y medidas mas allá de la inteligibilidad

La mayoría de los estudios identificados utilizan métricas objetivas de evaluación de señal como STOI, PESQ o la relación señal-ruido segmental, como indicadores primarios de rendimiento. Su correlación con la experiencia auditiva subjetiva y con el beneficio funcional en la vida diaria no siempre es directa: un sistema que maximiza STOI puede introducir artefactos audibles o generar una sonoridad no natural que comprometa la aceptabilidad clínica (Holube et al., 2020).

Schulz et al. (2026) y Ashkanichenarlogh et al. (2025) avanza en esta dirección al incluir evaluaciones perceptuales con oyentes reales. Saoji et al. (2024) añaden una perspectiva clínica al demostrar que la reducción de ruido basada en RNP modifica el rendimiento en tests audiológicos diagnósticos. Vaisberg et al. (2025) incorporan correlatos neurales del procesamiento auditivo como medidas convergentes del beneficio: en su muestra de 26 usuarios experimentados de audífono, el programa basado en AP se asoció con mayor precisión en escucha,

menor esfuerzo subjetivo y menor oxigenación de la corteza prefrontal izquierda en comparación con el programa estándar. Sin embargo, la evaluación ecológica en vida diaria sigue siendo escasa en el ámbito específico de la reducción de ruido basada en AP.

Qué se sabe. La evidencia disponible se distribuye en un gradiente que va desde evaluaciones objetivas en bancos de prueba hasta estudios con usuarios y medidas más cercanas al impacto rehabilitador. Las líneas emergentes fNIRS, candidatura a implante coclear, evaluación en usuarios con hipoacusia bajo condiciones diversas de ajuste, abren vías relevantes para conectar el beneficio técnico con la experiencia real.

Qué falta. Una evaluación más consistente con resultados que vayan más allá de la inteligibilidad: preferencia sostenida, fatiga o esfuerzo auditivo, aceptación a largo plazo. La evaluación ecológica y el seguimiento longitudinal son las brechas más urgentes.

Implicación clínica. En adaptación clínica, la adopción de funciones de AP en reducción de ruido debería acompañarse siempre de verificación funcional en ruido, con instrumentos que capturen no solo inteligibilidad sino comodidad, esfuerzo percibido y satisfacción sostenida.

Reconocimiento de entornos acústicos y automatización

El reconocimiento de entornos acústicos consiste en identificar automáticamente el tipo de situación sonora en la que se encuentra el usuario para ajustar en consecuencia el esquema de procesamiento del audífono. Este

componente opera como un módulo de control de nivel superior que habilita la conmutación automática de programas sin intervención del usuario, con el objetivo de reducir la carga de gestión del dispositivo y mejorar la coherencia del comportamiento del sistema (Hasemann & Krylova, 2024).

Un problema histórico para la investigación comparativa en este ámbito ha sido la limitada disponibilidad de conjuntos de datos abiertos representativos del uso en audífonos. Hüwel et al. (2020) presentaron HEAR-DS, un conjunto de datos binaural diseñado específicamente para el reconocimiento de entornos acústicos con orientación a aplicaciones en audífonos, facilitando comparaciones metodológicas reproducibles. Más recientemente, Zhong et al. (2025) propusieron AHEAD-DS, con etiquetado audiológicamente relevante, junto con un modelo de clasificación para audífonos denominado YAMNet+ en versiones iniciales y renombrado posteriormente como OpenYAMNet que adapta una arquitectura de propósito general a las 14 categorías pertinentes para el uso con audífonos. Esta línea evolutiva de HEAR-DS a AHEAD-DS evidencia la maduración de la infraestructura de investigación en reconocimiento de entornos orientado al audífono.

La automatización del procesamiento tiene implicaciones directas en la experiencia de usuario: una clasificación correcta y estable reduce la carga de gestión; por el contrario, clasificaciones erróneas o transiciones frecuentes e innecesarias pueden generar cambios de sonoridad percibidos como molestos y erosionar la confianza del usuario en el sistema

automático (Phonak, 2025; Sánchez et al., 2025).

Qué se sabe. La automatización basada en reconocimiento de escena es un caso de uso maduro en producto y conceptualmente sólido. La disponibilidad creciente de conjuntos de datos abiertos específicos como HEAR-DS o AHEAD-DS facilita la comparabilidad metodológica y desarrollo reproducible.

Qué falta. Conectar rendimiento del clasificador con resultados clínicos tangibles. La literatura es menos concluyente sobre cómo los errores y transiciones impactan satisfacción, confianza, control percibido y continuidad de uso.

Implicación clínica. En rehabilitación, el éxito del reconocimiento no se mide solo por “clasificar bien”, sino por comportarse bien: transiciones suaves, previsibilidad y posibilidad de control. Esto refuerza la necesidad de entrenamiento del usuario y verificación ecológica cuando se activan automatismos avanzados.

Del clasificador al resultado clínico

La precisión del clasificador de escena es una condición necesaria, pero no suficiente, para el beneficio clínico. La cadena causal que va desde el rendimiento técnico del clasificador hasta el impacto en la vida diaria del usuario implica múltiples eslabones: la clasificación correcta debe generar una selección de programa apropiada; el programa seleccionado debe producir una mejora perceptible en la señal; y esta mejora debe traducirse en mayor facilidad de comprensión, menor fatiga auditiva o mayor satisfacción para el

usuario concreto (Balling et al., 2021; Tasnim et al., 2024).

La literatura revisada presenta escasa evidencia que conecte directamente la mejora en precisión de clasificación con resultados funcionales medidos en usuarios reales. Los informes técnicos de fabricantes documentan tasas de clasificación correcta elevadas en condiciones de laboratorio, pero raramente incluyen evaluaciones con usuarios sobre el efecto de las conmutaciones automáticas en la experiencia de escucha cotidiana (Oticon, 2021; Phonak, 2025).

Qué se sabe. El rendimiento del clasificador es un requisito necesario pero insuficiente: el resultado rehabilitador depende de transiciones, previsibilidad, control percibido y coherencia del comportamiento automático.

Qué falta. Evidencia que conecte directamente precisión técnica con resultados clínicos en usuarios reales bajo condiciones representativas de uso cotidiano.

Implicación clínica. La adaptación de funciones automáticas debe acompañarse de educación al usuario sobre el comportamiento esperable del sistema y de mecanismos de control accesibles para situaciones en que la automatización no resulta óptima.

Personalización basada en preferencias

La personalización basada en preferencias del usuario constituye uno de los ejes más prometedores de la aplicación de IA en audífonos. Los enfoques de optimización bayesiana y comparaciones A/B permiten traducir percepciones subjetivas del usuario en ajustes cuantificables de los

parámetros del dispositivo, con un número limitado de iteraciones (Balling et al., 2021; Jensen et al., 2019; Tasnim et al., 2024).

La evidencia disponible procede de tres aproximaciones complementarias: factibilidad técnica, traslación a producto comercial y síntesis del campo. Jensen et al. (2019) demostraron la factibilidad de ajustar la ganancia del audífono mediante comparaciones A/B en 12 escenarios acústicos representativos con 20 participantes con hipoacusia, mostrando variabilidad interindividual significativa en las preferencias resultantes. Balling et al. (2021) describieron el sistema SoundSense Learn de Widex, que implementa optimización bayesiana sobre preferencias del usuario, presentando evidencia de su uso en condiciones de laboratorio y de campo. Tasnim et al. (2024) aportaron una revisión comprehensiva de los enfoques de AM aplicados a la personalización de audífonos, identificando los principales retos: definición de espacios de búsqueda relevantes, selección de escenas representativas para la evaluación y validación posterior del beneficio en uso cotidiano.

La personalización basada en preferencias plantea un cambio paradigmático respecto a los procedimientos de ajuste tradicionales: en lugar de partir de una prescripción derivada de la audiometría y aplicar ajustes finos según retroalimentación verbal del usuario, los sistemas de personalización por IA permiten al usuario explorar el espacio de configuraciones mediante interacciones simples y repetibles. Esto puede reducir la dependencia de la verbalización de

criterios acústicos, a menudo difícil para el usuario, y aumentar la sensibilidad a preferencias dependientes del contexto.

Qué se sabe. Los enfoques basados en comparaciones A/B y optimización bayesiana ofrecen un mecanismo factible para traducir percepciones subjetivas en ajustes cuantificables con un número limitado de iteraciones. Existe evidencia inicial de su aceptabilidad y utilidad en usuarios reales.

Qué falta. Validación a más largo plazo, en condiciones ecológicas, con seguimiento de la estabilidad de las preferencias derivadas y de su efecto en variables de uso cotidiano como satisfacción, adherencia, fatiga. Falta también estandarización de los protocolos clínicos: selección de escenas, número de comparaciones, criterios de parada.

Implicación clínica. La personalización basada en preferencias debe entenderse como un proceso integrado en la rehabilitación, no como un evento puntual. Su valor depende de su articulación con la educación del usuario, la verificación posterior y la posibilidad de revisión en función del contexto y la aclimatación.

Traslación industrial y ejemplos por marca

La traslación de la IA a producto comercial se documenta principalmente en informes técnicos de fabricantes. Estos documentos describen arquitecturas, decisiones de implementación y, en algunos casos, resultados de evaluaciones internas. En esta revisión se aplica un triple filtro de interpretación: (a) las descripciones de arquitectura e integración se registran como información descriptiva; (b) las afirmaciones de rendimiento se contrastan

con evidencia independiente cuando los diseños lo permiten; y (c) cualquier afirmación de beneficio se formula como hipótesis clínica en ausencia de validación externa.

Hasemann y Krylova (2024) documentan, en un informe técnico de Phonak, la incorporación de IA en la plataforma AutoSense OS, con énfasis en la clasificación de escenas y la adaptación automática del procesamiento; Phonak (2025) y Sánchez et al. (2025) describen la versión AutoSense OS 7.0 con foco en mayor precisión de adaptación a situaciones de escucha. Oticon (2021) presenta la plataforma More, que incorpora una red neuronal profunda entrenada con escenas sonoras reales para la representación de la señal antes de la amplificación, con la propuesta de “ofrecer al cerebro acceso al sonido completo”.

GN ReSound (2025a, 2025b) describe la función Intelligent Focus, que utiliza AP para mejorar el reconocimiento del habla en entornos ruidosos, posicionándola como complemento no sustituto del juicio del profesional clínico. Signia (2021) presenta la plataforma AX con énfasis en la separación binaural de fuentes mediante procesamiento dedicado por canal. Starkey, a través de Betlehem (2024), describe el uso de RNP para asistencia en reducción de ruido multibanda en su plataforma Edge AI. Widex (2021), por su parte, describe en su informe técnico la lógica de aprendizaje de preferencias del usuario mediante optimización bayesiana, sobre la que se construye su sistema SoundSense Learn.

Las descripciones técnicas de cada fabricante difieren en nivel de detalle y en

la naturaleza de la evidencia interna reportada. Los informes habitualmente incluyen métricas de rendimiento técnico como tasas de clasificación correcta, mejoras en STOI o equivalentes, pero raramente reportan evaluaciones con usuarios independientes en condiciones ecológicas. Esta característica de la documentación técnica de fabricantes refuerza la necesidad de mantener separados los niveles de evidencia en la síntesis: las afirmaciones documentales de los fabricantes son útiles para describir el panorama industrial, pero no equivalen a validación científica independiente. La relación completa de documentos técnicos consultados se incluye en el Anexo B.

IA, sensores y funciones de salud

Los audífonos contemporáneos han ido incorporando progresivamente sensores adicionales como acelerómetros, giroscopios, sensores de proximidad y conectividad con sistemas de salud digital (Fabry & Bhowmik, 2021). Esta ampliación de capacidades permite explorar funciones que trascienden la rehabilitación auditiva clásica: detección de caídas, seguimiento de actividad física, alerta de eventos relevantes y monitorización pasiva del comportamiento auditivo.

La integración de IA con datos multimodales como audio, movimiento, comportamiento de uso, abre vías de investigación sobre la relación entre la experiencia auditiva y el bienestar general. Los sistemas de IA pueden agregar información de múltiples fuentes para identificar patrones, p. ej., cambios en el uso del dispositivo, alteraciones en los entornos frecuentados, modificaciones en

el patrón de actividad física, que pueden señalar necesidades de ajuste, problemas técnicos o cambios en el estado del usuario que requieren atención clínica.

Esta ampliación del rol del audífono plantea oportunidades, p. ej., seguimiento longitudinal más rico, posibilidad de detección temprana de problemas, pero también desafíos relevantes en términos de privacidad, consentimiento informado, integración en flujos clínicos y equidad en el acceso (Nada & Moaty, 2025).

DISCUSIÓN

Los hallazgos de esta revisión permiten organizar la discusión en torno a cuatro ejes complementarios: las implicaciones clínicas para la adaptación y el ajuste fino del audífono; las herramientas de evaluación centradas en la persona necesarias para capturar el beneficio rehabilitador; la integración con teleaudiología y evaluación ecológica; y las consideraciones éticas relativas a privacidad, sesgos y explicabilidad.

Implicaciones clínicas en adaptación y ajuste fino

La incorporación de IA en audífonos modifica el flujo clínico de adaptación en varios puntos. La automatización del reconocimiento de entornos reduce la necesidad de manejar manualmente programas, lo cual puede simplificar la experiencia del usuario, pero también plantea la cuestión de cómo educar sobre el comportamiento del sistema y cómo proporcionar mecanismos de control accesibles para situaciones en que la automatización no resulta óptima (Phonak, 2025).

La personalización basada en preferencias del usuario abre la posibilidad de un ajuste fino más sensible al contexto y a la subjetividad individual, complementando la prescripción derivada de la audiometría con información de uso real (Balling et al., 2021; Jensen et al., 2019). Su valor clínico depende, sin embargo, de la integración en protocolos estructurados que incluyan selección de escenas representativas, número limitado de comparaciones por sesión y verificación posterior del beneficio en condiciones ecológicas.

Evaluación centrada en la persona

La evaluación del beneficio rehabilitador en audífonos con funciones de IA requiere instrumentos sensibles a dimensiones que van más allá de la inteligibilidad: satisfacción, esfuerzo de escucha, fatiga auditiva, adherencia y participación social. Cuestionarios validados como el SSQ (Speech, Spatial and Qualities of Hearing Scale; Gatehouse & Noble, 2004) y el APHAB (Abbreviated Profile of Hearing Aid Benefit; Cox & Alexander, 1995) ofrecen marcos de referencia consolidados para capturar estas dimensiones.

Para capturar la variabilidad situacional de la experiencia auditiva, la EMA se ha establecido como el método de referencia en investigación audiológica (Holube et al., 2020). Schinkel-Bielefeld et al. (2024) ofrecen una guía práctica sobre su implementación en atención audiológica. Vercammen et al. (2023) demostraron que la combinación de EMA autodesencadenada con análisis de lenguaje natural puede capturar información ecológicamente válida sobre la experiencia de uso del audífono en la

vida diaria, con menor sesgo retrospectivo que los cuestionarios convencionales.

Teleaudiología, ajuste remoto y evaluación ecológica

La expansión de la teleaudiología como modalidad de atención ha sido uno de los desarrollos más significativos de los últimos años en el campo de la rehabilitación auditiva. Çelikgün y Büyükkal (2023) evaluaron el efecto de la tecnología de ajuste remoto sobre la satisfacción con el audífono, encontrando niveles de satisfacción comparables a los del seguimiento presencial, con la ventaja añadida de mayor conveniencia y menores barreras de acceso.

La IA facilita la teleaudiología en varias dimensiones complementarias: los sistemas de aprendizaje de preferencias generan un flujo de información ecológicamente válido revisable a distancia; los clasificadores de escena registran datos pasivos sobre el comportamiento del sistema en la vida diaria que pueden consultarse a distancia. La integración entre EMA y teleaudiología puede operar como un circuito de retroalimentación: el usuario reporta dificultad en contexto, el profesional ajusta presencial o remotamente y el usuario reevalúa en escenas comparables. Sin embargo, la implementación de teleaudiología con componentes de IA y EMA plantea desafíos de equidad y acceso que no deben obviarse, particularmente en personas de edad avanzada con menor competencia tecnológica.

Privacidad, sesgos y explicabilidad

La integración de la IA en audífonos implica la recogida, transmisión y

procesamiento de datos de naturaleza sensible. Los audífonos modernos conectados capturan audio ambiental de forma continua, registran datos de comportamiento de uso y, en plataformas con sensores integrados, pueden recoger indicadores de salud fisiológicos (Fabry & Bhowmik, 2021). Esta capacidad de monitorización permanente requiere marcos de gobernanza claros, consentimiento informado específico y medidas de seguridad técnica robustas.

Los modelos de AM y AP están condicionados por los datos con los que fueron entrenados. Si los conjuntos de entrenamiento no representan adecuadamente la diversidad de perfiles audiométricos, edades, entornos culturales e idiomas de los usuarios, los modelos pueden exhibir sesgos sistemáticos que perjudican desproporcionadamente a determinados subgrupos. La literatura revisada raramente aborda la representatividad de los datos de entrenamiento, lo que representa una limitación notable desde la perspectiva de la equidad en salud (Nada & Moaty, 2025).

La “explicabilidad” clínica no implica revelar detalles de arquitectura, sino asegurar que el usuario comprende el comportamiento del dispositivo: cuándo actúa, qué cambia, cómo recuperar control; y que el profesional dispone de criterios para ajustar o desactivar funciones cuando comprometen confort, confianza o adherencia.

CONCLUSIONES

Esta revisión ha sintetizado la evidencia disponible sobre aplicaciones de IA en audífonos en cuatro ejes temáticos:

reducción de ruido y realce del habla, reconocimiento de entornos acústicos y automatización, personalización basada en preferencias del usuario, y soporte remoto con evaluación ecológica.

Beneficio condicionado por restricciones embebidas. En reducción de ruido basada en AP, la plausibilidad técnica no depende solo del rendimiento de separación, sino de su compatibilidad con operación causal, baja latencia, consumo y estabilidad perceptiva. El beneficio reportado debe interpretarse siempre como dependiente del escenario acústico, del perfil del usuario y del criterio de evaluación (Diehl et al., 2023; Schulz et al., 2026; Vaisberg et al., 2025).

De “clasificar escenas” a “mejorar experiencia”. En reconocimiento de entornos, el rendimiento del clasificador es un requisito necesario pero insuficiente: el resultado rehabilitador depende de transiciones, previsibilidad, control percibido y coherencia del comportamiento automático (Hüwel et al., 2020; Zhong et al., 2025).

La personalización por preferencias es clínicamente prometedora si se diseña como proceso. Los enfoques basados en comparaciones A/B y optimización bayesiana ofrecen un mecanismo para traducir percepciones subjetivas en ajustes cuantificables con un número limitado de iteraciones. Su adopción clínica depende de la eficiencia del protocolo, la selección de escenas relevantes y la validación posterior en vida diaria (Balling et al., 2021; Jensen et al., 2019).

La evidencia abierta y los informes técnicos de fabricantes cumplen funciones distintas. La literatura de

acceso abierto aporta métodos, comparadores y evaluaciones iniciales; los informes técnicos de fabricantes describen traslación industrial, supuestos y arquitectura, pero no deben equipararse a evidencia revisada por pares.

Criterio rehabilitador: impacto centrado en la persona. El valor clínico de la IA no se determina únicamente por métricas objetivas, sino por su efecto en resultados centrados en la persona como satisfacción, esfuerzo o fatiga, participación y uso continuado, y por su integración en protocolos de adaptación, educación y seguimiento.

Cierre vinculado a objetivos

En relación con el Objetivo 1, la revisión permitió describir y comparar enfoques de IA para reducción de ruido y realce del habla enfatizando las restricciones embebidas que condicionan su viabilidad (causalidad, latencia, cómputo y estabilidad perceptiva), evidenciando que el beneficio técnico reportado debe interpretarse como dependiente del escenario y del criterio de evaluación.

Respecto al Objetivo 2, se sintetizó el papel del reconocimiento de entornos acústicos como mecanismo de automatización, destacando que la relevancia clínica no se agota en la precisión de clasificación, sino que depende de transiciones, coherencia de comportamiento y control percibido, e incluyendo la línea evolutiva de conjuntos de datos abiertos específicos para el dominio audiológico (HEAR-DS, AHEAD-DS).

En el Objetivo 3, se integraron métodos de personalización basados en preferencias,

mostrando que las comparaciones A/B y estrategias de optimización pueden traducir juicios subjetivos en ajustes accionables, aunque su adopción clínica queda condicionada por la carga de interacción, la dependencia del contexto y la necesidad de validación ecológica.

Para el Objetivo 4, se caracterizó la traslación industrial mediante documentación técnica pública aplicando un triple filtro de interpretación que establece una separación explícita entre afirmaciones de rendimiento, validación interna y evidencia revisada por pares.

Finalmente, se discutieron implicaciones clínicas y organizativas subrayando que el impacto rehabilitador de la IA requiere protocolos centrados en la persona, integración con teleaudiología, herramientas de evaluación ecológica y prácticas de transparencia, privacidad y cautela en la generalización.

En síntesis, este trabajo contribuye a integrar la evidencia técnica y clínica en un marco interpretativo orientado a la práctica audiológica real, destacando la necesidad de enfoques centrados en la persona para la implementación efectiva de tecnologías basadas en inteligencia artificial.

Limitaciones

Las revisiones recientes coinciden en una misma observación: el desarrollo técnico avanza más deprisa que su validación clínica (Andersen et al., 2021; Nada & Moaty, 2025; Tasnim et al., 2024). Es una brecha que reaparece a lo largo del trabajo: en la falta de evaluación con usuarios reales en reducción de ruido, y en la cautela con que deben interpretarse los resultados que no se han probado en

hardware real ni tienen en cuenta las restricciones del dispositivo. Por eso enmarca tanto las limitaciones como las líneas futuras.

La presente revisión presenta varias limitaciones que conviene reconocer.

En primer lugar, restringir la búsqueda a documentación de acceso abierto puede haber introducido un sesgo de disponibilidad: estudios relevantes publicados en revistas con acceso restringido o tesis no indexadas quedan fuera del análisis.

En segundo lugar, el enfoque narrativo —aunque metodológicamente justificado por la heterogeneidad del campo— no permite estimaciones agregadas comparables a las de una revisión sistemática y, en consecuencia, las conclusiones tienen carácter interpretativo.

En tercer lugar, una parte sustancial de la traslación industrial se documenta exclusivamente a través de informes técnicos de fabricantes que, pese a tratarse con el triple filtro descrito en el apartado de Traslación industrial y ejemplos por marca, no equivalen a validación científica independiente y limitan la generalización de las afirmaciones sobre rendimiento.

En cuarto lugar, la velocidad de evolución del campo —con publicaciones relevantes apareciendo durante el propio proceso de redacción del trabajo— hace que cualquier síntesis tenga necesariamente un carácter provisional.

Estas limitaciones definen las condiciones en que deben leerse los resultados y hasta qué punto pueden generalizarse, más que cuestionar su validez.

Prospectiva

Ensayos pragmáticos centrados en la persona. Diseñar estudios con usuarios de audífono que incluyan resultados primarios funcionales: preferencia sostenida, satisfacción, esfuerzo o fatiga, y secundarios técnicos: rendimiento en habla en ruido, con seguimiento suficiente para capturar aclimatación y estabilidad del beneficio a largo plazo.

Evaluación ecológica integrada en el protocolo. Implementar sistemas EMA de baja carga — mediante microcuestionarios contextualizados — y vincularlos a registros automáticos de escena y de uso del dispositivo, con el fin de evaluar la automatización y la personalización en las condiciones reales en que se aplican. Establecer, a partir de esos datos, reglas de decisión clínica sobre cuándo modificar el ajuste, cuándo desactivar funciones automáticas y cuándo reiniciar el proceso de aprendizaje de preferencias.

Benchmarks abiertos representativos del dispositivo. Promover conjuntos de datos y protocolos comparables para habla en ruido y clasificación de escenas que utilicen señales y condiciones cercanas a audífonos como binauralidad, micrófonos, ruido real, escenas mixtas, permitiendo comparaciones reproducibles entre enfoques.

Modelos eficientes e híbridos AP + procesamiento de señal clásico. Priorizar arquitecturas compactas, cuantización y estrategias híbridas que combinen procesamiento clásico: robusto y eficiente, con módulos de AP donde aporten valor incremental, manteniendo latencia baja y estabilidad perceptiva para uso continuo.

Protocolos clínicos para personalización por preferencias. Estandarizar procedimientos: selección de escenas, número máximo de comparaciones por sesión, instrucciones al usuario, criterios de parada y verificación posterior. Evaluar explícitamente la interacción entre preferencia, contexto y aclimatación para evitar ajustes frágiles o inestables.

Gobernanza, privacidad y auditoría de sesgos. Establecer buenas prácticas de consentimiento, minimización de datos y trazabilidad, junto con estrategias de auditoría de desempeño por idioma, entorno y perfiles de usuario. Traducir estas prácticas a documentación clínica: qué funciones están activas, qué datos se usan y qué limitaciones se comunican al usuario.

REFERENCIAS

- Andersen, A. H., Santurette, S., Pedersen, M. S., Alickovic, E., Fiedler, L., Jensen, J., & Behrens, T. (2021). Creating clarity in noisy environments by using deep learning in hearing aids. *Seminars in Hearing*, 42(3), 260–281. <https://doi.org/10.1055/s-0041-1735134>
- Ashkanichenarlogh, V., Folkeard, P., Scollie, S., Kühnel, V., & Parsa, V. (2025). Objective evaluation of a deep learning-based noise reduction algorithm for hearing aids under diverse fitting and listening conditions. *Trends in Hearing*, 29. <https://doi.org/10.1177/23312165251396644>
- Aubreville, M., Ehrensperger, K., Rosenkranz, T., Graf, B., Puder, H., & Maier, A. (2018). Deep denoising for hearing aid applications. En 2018 16th *International Workshop on Acoustic Signal Enhancement (IWAENC)* (pp. 361–365). IEEE. <https://doi.org/10.1109/IWAENC.2018.8521369>
- Balling, L. W., Mølgaard, L. L., Townend, O., & Nielsen, J. B. B. (2021). The collaboration between hearing aid users and artificial intelligence to optimize sound. *Seminars in Hearing*, 42(3), 282–294. <https://doi.org/10.1055/s-0041-1735135>
- Betlehem, T. (2024). *Leveraging DNN in Starkey Edge AI (DNN assisted MNR)* [Informe técnico de fabricante]. Starkey.
- Cabrera, A., del Río, S., Leiva, J. C., Pérez, A., & Ferrer, L. G. (2024). La inteligencia artificial al servicio de la ayuda auditiva. *Medicentro Electrónica*, 28, e4153. <https://medicentro.sld.cu/index.php/medicentro/article/view/4153>
- Çelikgün, B., & Büyükkal, F. (2023). The effect of remote fitting technology on hearing aid satisfaction. *The Egyptian Journal of Otolaryngology*, 39(1), 83. <https://doi.org/10.1186/s43163-023-00447-7>
- Cox, R. M., & Alexander, G. C. (1995). *The Abbreviated Profile of Hearing Aid Benefit*. *Ear and Hearing*, 16(2), 176–186.

- <https://doi.org/10.1097/00003446-199504000-00005>
- Diehl, P. U., Zilly, H., Sattler, F., Singer, Y., Kepp, K., Berry, M., Hasemann, H., Zippel, M., Kaya, M., Meyer-Rachner, P., Pudzuhn, A., Hofmann, V. M., Vormann, M., & Sprengel, E. (2023). Deep learning-based denoising streamed from mobile phones improves speech-in-noise understanding for hearing aid users. *Frontiers in Medical Engineering*, 1, 1281904. <https://doi.org/10.3389/fmede.2023.1281904>
- Fabry, D. A., & Bhowmik, A. K. (2021). Improving speech understanding and monitoring health with hearing aids using artificial intelligence and embedded sensors. *Seminars in Hearing*, 42(3), 295–308. <https://doi.org/10.1055/s-0041-1735136>
- Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, 43(2), 85–99. <https://doi.org/10.1080/14992020400050014>
- GN ReSound. (2025a). *Intelligent Focus: How deep learning augments human intelligence* [Informe técnico de fabricante]. GN ReSound. Consultado el 15 de enero de 2026 en <https://pro.resound.com/en-us/research/features-explained>
- GN ReSound. (2025b). *Features explained: Intelligent Focus* [Informe técnico de fabricante].
- GN ReSound. Consultado el 15 de enero de 2026 en <https://pro.resound.com/en-us/research/features-explained>
- Hasemann, H., & Krylova, A. (2024). *Artificial intelligence in hearing aid technology* [Phonak Insight]. Phonak. Consultado el 15 de enero de 2026 en https://www.phonak.com/content/dam/phonak/en/evidence-library/white-paper/technical-paper/PH_Insight_ArtificialIntelligenceInHearingAidTechnology.pdf
- Holube, I., von Gablenz, P., & Bitzer, J. (2020). Ecological momentary assessment in hearing research: Current state, challenges, and future directions. *Ear and Hearing*, 41(Suppl 1), 79S–90S. <https://doi.org/10.1097/AUD.0000000000000934>
- Hüwel, A., Adiloğlu, K., & Bach, J.-H. (2020). Hearing aid research data set for acoustic environment recognition. En *ICASSP 2020 — 2020 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 706–710). IEEE. <https://doi.org/10.1109/ICASSP40776.2020.9053611>
- Jensen, N. S., Hau, O., Nielsen, J. B. B., Nielsen, T. B., & Legarth, S. V. (2019). Perceptual effects of adjusting hearing-aid gain by means of a machine-learning approach based on individual user preference. *Trends in Hearing*, 23, 2331216519847413.

- <https://doi.org/10.1177/2331216519847413>
- Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbaek, G., Alladi, S., Ames, D., Banerjee, S., Burns, A., Brayne, C., Fox, N. C., Ferri, C. P., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Nakasujja, N., Rockwood, K., ... Mukadam, N. (2024). Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*, 404(10452), 572–628. [https://doi.org/10.1016/S0140-6736\(24\)01296-0](https://doi.org/10.1016/S0140-6736(24)01296-0)
- Nada, N., & Moaty, A. S. (2025). Beyond amplification: How artificial intelligence is redefining hearing. *The Egyptian Journal of Otolaryngology*, 41, artículo 211. <https://doi.org/10.1186/s43163-025-00965-6>
- Oticon. (2021). *Oticon More* [Informe técnico de fabricante]. Oticon. Consultado el 15 de enero de 2026 en https://www.hnrspeechandhearing.com/wp-content/uploads/2021/07/Oticon-More_HNR.pdf
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Phonak. (2025). *Ficha técnica de AM: clasificador de sonido AutoSense OSTM* [Ficha técnica]. Phonak. Consultado el 15 de enero de 2026 en https://www.phonak.com/content/dam/celum/phonak/local-assets/ca/documents/PH_FactSheet_ML-AutoSenseOS_210x297_EN.pdf
- Sánchez, C., Giurda, R., Hobi, S., & Preuss, M. (2025). *AutoSense OS 7.0 adapts more precisely to listening situations* [Informe técnico de fabricante]. Phonak. Consultado el 15 de enero de 2026 en <https://audiologyblog.phonakpro.com/phonak-infinio-ultra-adapts-in-real-time-for-high-quality-sound-in-every-situation/>
- Saoji, A. A., Sheikh, B. A., Bertsch, N. J., Goulson, K. R., Graham, M. K., McDonald, E. A., Bross, A. E., Vaisberg, J. M., Kühnel, V., Voss, S. C., Qian, J., Hogan, C. H., & DeJong, M. D. (2024). How does deep neural network-based noise reduction in hearing aids impact cochlear implant candidacy? *Audiology Research*, 14(6), 1114–1125. <https://doi.org/10.3390/audiolres14060092>
- Schinkel-Bielefeld, N., Burke, L., Holube, I., Iankilevitch, M., Jenstad, L. M.,

- Lelic, D., Naylor, G., Singh, G., Smeds, K., von Gablenz, P., Wolters, F., & Wu, Y.-H. (2024). Implementing ecological momentary assessment in audiological research: Opportunities and challenges. *American Journal of Audiology*, 33(3), 648–673. https://doi.org/10.1044/2024_AJA-23-00249
- Schröter, H., Rosenkranz, T., Escalante-B., A. N., Aubreville, M., & Maier, A. (2020). CLCNet: Deep learning-based noise reduction for hearing aids using complex linear coding. En *ICASSP 2020 — 2020 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 6691–6695). IEEE. <https://doi.org/10.1109/ICASSP40776.2020.9053563>
- Schulz, K. M. B., Wittmann, H., Hackenberg, L., Jürgens, T., & Tchorz, J. (2026). Effects of low-latency deep-learning-based noise reduction on speech intelligibility for normal hearing, hearing-impaired listeners and cochlear implant users. *Frontiers in Audiology and Otology*, 3, 1746635. <https://doi.org/10.3389/fauot.2025.1746635>
- Signia. (2021). *Signia AX platform brochure* [Folleto técnico de fabricante]. Signia. Consultado el 15 de enero de 2026 en https://www.signia-pro.com/-/media/signia/global/files/signia-ax/signia-ax-platform-brochure_2021-02.pdf
- Stone, M. A., & Moore, B. C. J. (2005). Tolerable hearing aid delays. IV. Effects on subjective disturbance during speech production by hearing-impaired subjects. *Ear and Hearing*, 26(2), 225–235. <https://doi.org/10.1097/00003446-200504000-00009>
- Stone, M. A., Moore, B. C. J., Meisenbacher, K., & Derleth, R. P. (2008). Tolerable hearing aid delays. V. Estimation of limits for open canal fittings. *Ear and Hearing*, 29(4), 601–617. <https://doi.org/10.1097/AUD.0b013e3181734ef2>
- Tasnim, N. Z., Ni, A., Lobarinas, E., & Kehtarnavaz, N. (2024). A review of machine learning approaches for the personalization of amplification in hearing aids. *Sensors*, 24(5), 1546. <https://doi.org/10.3390/s24051546>
- Vaisberg, J. M., Dang, C., Jiang, Y., Qian, J., & Russo, F. A. (2025). Brain benefits of deep learning-based noise management in experienced hearing aid users using functional near infrared spectroscopy. *Scientific Reports*, 15, 41815. <https://doi.org/10.1038/s41598-025-25801-y>
- Vercammen, C., Oosthuizen, I., Manchaiah, V., Ratinaud, P., Launer, S., & Swanepoel, D. W. (2023). Real-life and real-time hearing aid experiences: Insights from self-initiated ecological momentary assessments and

natural language analysis.
Frontiers in Digital Health, 5,
1104308.

<https://doi.org/10.3389/fdgth.2023.1104308>

Widex. (2021). *The collaboration between hearing aid users and artificial intelligence to optimize sound* [Informe técnico de fabricante]. Widex. Consultado el 15 de enero de 2026 en https://www.widexpro.com/-/media/widex/global/files/evidence/Widex-evidence_2021-09_The-Collaboration-between-Hearing-Aid-Users-and-Artificial-Intelligence.pdf

World Health Organization. (2021). World report on hearing. WHO. <https://www.who.int/publications/i/item/world-report-on-hearing>

World Health Organization. (2024). Deafness and hearing loss [Fact sheet]. WHO. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>

Zhong, H., Buchholz, J. M., Maclaren, J., Carlile, S., & Lyon, R. (2025). *A dataset and model for recognition of audiologically relevant environments for hearing aids: AHEAD-DS and YAMNet+* [Preprint]. arXiv. <https://arxiv.org/abs/2508.10360>

Anexo A. Estudios y fuentes principales incluidos

La Tabla 6 resume los estudios clave incluidos en la revisión con texto completo de acceso abierto, ordenados por eje temático.

Tabla 6.

Selección de estudios clave incluidos en la revisión (acceso abierto).

Referencia	Tipo	IA	Datos / contexto	Resultado principal	Relevancia clínica
Aubreville et al. (2018)	Artículo (IWAENC/IE EE)	AP — ganancia tipo Wiener	Base de ruidos reales; enfoque baja latencia	Mejora sobre referencia en métricas y pruebas subjetivas	Viabilidad de AP con restricciones de audífonos
Schröter et al. (2020)	Artículo (ICASSP/IEE E)	AP — CLCNet	Incluye ruido grabado con audífonos; diseño causal	Mejoras en condiciones ruidosas; diseño compatible con tiempo real	Importancia de datos del dispositivo real
Andersen et al. (2021)	Artículo (Seminars in Hearing)	AP — revisión técnica integradora	Síntesis del estado del arte; ejemplos en oyentes con hipoacusia	Marco de cómo las RNP generan claridad en entornos ruidosos; compromisos entre métricas objetivas, latencia y valoración	Referencia conceptual para evaluar AP en reducción de ruido en audífonos
Diehl et al. (2023)	Artículo (Frontiers Med Eng)	AP — denoising en teléfono inteligente	Tiempo real en móvil; usuarios con hipoacusia	Mejora 1,6 dB en SRT; aumento >40 % en valoraciones	Traslación vía

Referencia	Tipo	IA	Datos / contexto	Resultado principal	Relevancia clínica
			moderada-severa	subjetivas; latencia total 65–75 ms	ecosistema móvil
Saoji et al. (2024)	Artículo (Audiology Research, MDPI)	AP — RNP en audífonos comerciales	10 pacientes con hipoacusia severa	Mejora 20–32 puntos porcentuales sobre programa basal; impacto en candidatura a IC	Interacción IA con decisiones clínicas
Ashkanichenarlogh et al. (2025)	Artículo (Trends in Hearing)	AP — evaluación objetiva	Grabaciones con KEMAR; múltiples ajustes (1.152 grabaciones)	Estimación HASPI bajo condiciones diversas	Comparabilidad y robustez por condiciones
Schulz et al. (2026)	Artículo (Frontiers Audiol Otol)	AP — Conv-TasNet, latencia 1 ms	NH (n=20), HI (n=19), IC (n=10 analizados de 12 inscritos); OLSA en olnoise	Mediana SRT: NH -0,9 dB; HI 0,8 dB; IC 5,7 dB	Evaluación bajo restricciones clínicas reales
Vaisberg et al. (2025)	Artículo (Scientific Reports)	AP — gestión de ruido	Medición fNIRS de oxigenación prefrontal en 26 usuarios experimentados	Mayor precisión, menor esfuerzo subjetivo y menor oxigenación	Relevancia para esfuerzo o fatiga auditiva

Referencia	Tipo	IA	Datos / contexto	Resultado principal	Relevancia clínica
				prefrontal con programa AP	
Hüwel et al. (2020)	Actas (ICASSP)	Dataset HEAR-DS	Binaural; reconocimiento de entorno	Dataset abierto para entrenamiento y evaluación	Facilita benchmarks reproducibles
Zhong et al. (2025)	Preprint (arXiv)	AP — AHEAD-DS + YAMNet+/OpenYA MNet	Dataset abierto; despliegue edge	Referencia con buenas métricas en 14 categorías	Automatización contextual reproducible
Jensen et al. (2019)	Artículo (Trends in Hearing)	AM — ajuste de ganancia por preferencias	Comparaciones A/B en 12 escenarios; 20 participantes con hipoacusia	Factibilidad; variabilidad interindividual	Apoyo a ajuste fino centrado en persona
Balling et al. (2021)	Artículo (Seminars in Hearing)	AM — optimización bayesiana (SoundSense Learn)	Entrada continua del usuario; lab. y campo	Resumen de evidencia	Personalización entre visitas
Tasnim et al. (2024)	Revisión (Sensors)	AM — personalización	Revisión de enfoques y retos	Síntesis y estrategias prometedoras	Marco para adopción clínica
Vercammen et al. (2023)	Artículo (Frontiers Digit Health)	EMA + análisis de lenguaje natural	EMA autodesencadena da; usuarios reales	Captura de experiencia contextual en vida diaria	Evaluación ecológica en rehabilitación
Schinkel-Bielefeld et al. (2024)	Artículo (Am. Journal of Audiology)	EMA — implementación	Guía práctica para atención audiológica	Recomendaciones de diseño y análisis	Integración de EMA en seguimiento clínico

Referencia	Tipo	IA	Datos / contexto	Resultado principal	Relevancia clínica
Fabry & Bhowmik (2021)	Artículo (Seminars in Hearing)	IA + sensores embebidos	Integración con detección actividad y salud	Visión de tendencias e implicaciones	Extiende rehabilitación a salud digital
Cabrera et al. (2024)	Revisión (SciELO Cuba)	IA en ayuda auditiva (español)	Perspectiva clínica y divulgativa	Síntesis de aplicaciones	Accesible para contexto hispanohablante
Nada & Moaty (2025)	Revisión (Egyptian J Otolaryngol)	IA en audífonos — panorámica	Revisión comprehensiva	Síntesis de 4 ejes; brecha evidencia/producto	Marco general e implicaciones clínicas
Çelikgün & Büyükkal (2023)	Artículo (Egyptian J Otolaryngol)	Teleaudiología — ajuste remoto	Comparación presencial vs. remoto	Satisfacción comparable	Factibilidad del seguimiento a distancia

Nota. AM = aprendizaje automático; AP = aprendizaje profundo; EMA = evaluación momentánea ecológica; NR = reducción de ruido (noise reduction); SRT = umbral de recepción del habla (speech reception threshold); SNR = relación señal-ruido (signal-to-noise ratio).

Anexo B. Documentación técnica de fabricantes consultada

La Tabla 7 recoge todos los documentos técnicos de fabricantes utilizados en la revisión, con sus datos de identificación y fecha de consulta.

Tabla 7.

Documentos técnicos de fabricantes incluidos en la revisión.

Fabricante / autoría	Título del documento	Tipo	Año	Fecha de consulta
Phonak (Hasemann y Krylova)	Artificial intelligence in hearing aid technology	Documento técnico	2024	Enero 2026
Phonak	ML Fact Sheet: Sound Classifier AutoSense OSTM	Ficha técnica	2025	Enero 2026
Phonak (Sánchez et al.)	AutoSense OS 7.0 adapts more precisely to listening situations	Informe técnico	2025	Enero 2026
Oticon	Oticon More	Informe técnico	2021	Enero 2026
GN ReSound	Intelligent Focus: How deep learning augments human intelligence	Informe técnico	2025	Enero 2026
GN ReSound	Features explained: Intelligent Focus	Informe técnico	2025	Enero 2026
Signia	Signia AX platform brochure	Folleto técnico	2021	Enero 2026
Starkey (Betlehem)	Leveraging DNN in Starkey Edge AI (DNN assisted MNR)	Informe técnico	2024	Enero 2026
Widex	The collaboration between hearing aid users and artificial intelligence to optimize sound	Informe técnico	2021	Enero 2026

Nota. Todos los documentos fueron consultados durante enero de 2026. Las URL completas se proporcionan en la sección Referencias.

Anexo C. Glosario de términos técnicos y acrónimos

La Tabla 8 recoge los principales términos técnicos y acrónimos empleados en este trabajo. Las definiciones son operativas y responden al uso concreto de cada término en el contexto de la revisión.

Tabla 8.

Glosario de términos técnicos y acrónimos empleados en la revisión.

Término / acrónimo	Definición operativa en este trabajo
AM (aprendizaje automático)	Rama de la IA que emplea modelos entrenados con descriptores acústicos definidos de forma explícita para abordar tareas de clasificación, discriminación o decisión. Equivale al término inglés machine learning (ML).
AP (aprendizaje profundo)	Subcampo del AM que utiliza arquitecturas neuronales con múltiples capas ocultas para aprender representaciones jerárquicas a partir de grandes volúmenes de datos. Equivale al término inglés deep learning (DL).
APHAB (<i>Abbreviated Profile of Hearing Aid Benefit</i>)	Cuestionario autoadministrado de 24 ítems que mide el beneficio percibido del audífono en situaciones cotidianas (Cox & Alexander, 1995). Evalúa cuatro subescalas: facilidad de comunicación, reverberación, entornos ruidosos y aversividad.
A/B (comparación emparejada)	Paradigma de evaluación en el que el usuario elige entre dos configuraciones sonoras sin verbalizar criterios acústicos. Base de la personalización por preferencias
Causalidad (procesamiento causal)	Capacidad de un algoritmo de operar sin acceder a tramas futuras de la señal. Requisito imprescindible para el procesamiento en tiempo real en audífonos
DNN (<i>Deep Neural Network</i>)	Véase RNP.
EMA (evaluación momentánea ecológica)	Metodología de recogida de datos que captura experiencias, comportamientos o estados del usuario en el momento en que se producen en la vida diaria, minimizando el sesgo retrospectivo. Equivale al término inglés ecological momentary assessment.

Término / acrónimo	Definición operativa en este trabajo
fNIRS (espectroscopia de infrarrojo cercano funcional)	Técnica de neuroimagen no invasiva que mide cambios en la oxigenación de la corteza cerebral mediante luz infrarroja cercana. Se emplea en investigación auditiva como marcador del esfuerzo cognitivo de escucha.
HASPI (<i>Hearing Aid Speech Perception Index</i>)	Índice de predicción de inteligibilidad del habla auditivamente informado, diseñado para estimar el rendimiento perceptual en oyentes con hipoacusia con mayor precisión que métricas estándar como STOI.
HEAR-DS (<i>Hearing Aid Research Data Set</i>)	Conjunto de datos binaural para el reconocimiento de entornos acústicos orientado a aplicaciones en audífonos, presentado por Hüwel et al. (2020). Fue uno de los primeros conjuntos abiertos específicamente diseñados para este dominio.
IA (inteligencia artificial)	Conjunto de técnicas computacionales para inferir estados u optimizar decisiones a partir de datos. En este trabajo engloba AM, AP y sus combinaciones
Latencia	Retardo temporal entre la entrada acústica al micrófono y la salida procesada al oído. El límite tolerable se sitúa en torno a 5–6 ms en adaptaciones abiertas (Stone et al., 2008) y 20–30 ms en cerradas (Stone & Moore, 2005); en AP de baja latencia son alcanzables valores de ~1 ms (Schulz et al., 2026).
NR (<i>Noise Reduction</i> / reducción de ruido)	Familia de algoritmos que atenúan las componentes de señal identificadas como ruido para mejorar la relación señal-ruido percibida. Se distingue entre NR clásica (estadística) y NR basada en AP.
Optimización bayesiana	Método de optimización basado en un modelo probabilístico actualizable que minimiza el número de evaluaciones necesarias para encontrar el óptimo. En audífonos, reduce la carga de comparaciones del usuario en sistemas de personalización
PESQ (<i>Perceptual Evaluation of Speech Quality</i>)	Métrica objetiva normalizada para evaluar la calidad de señales de voz procesadas, correlacionada con valoraciones subjetivas de calidad percibida. Ampliamente usada en la evaluación de algoritmos de reducción de ruido.

Término / acrónimo	Definición operativa en este trabajo
RNP (Red Neuronal Profunda)	Arquitectura de AP con múltiples capas de neuronas que aprende representaciones jerárquicas. Equivale a DNN. Se emplea en audífonos para estimación de máscaras o ganancias.
SNR (<i>Signal-to-Noise Ratio</i> / relación señal-ruido)	Medida de la diferencia de nivel entre la señal deseada (habla) y el ruido de fondo, expresada en decibelios. Un SNR positivo indica que el habla supera en nivel al ruido.
SSQ (<i>Speech, Spatial and Qualities of Hearing Scale</i>)	Escala de Habla, Espacial y Cualidades Auditivas. Cuestionario autoadministrado de 49 ítems diseñado para evaluar el funcionamiento auditivo en tres dominios: comprensión del habla en distintos escenarios, localización espacial del sonido y cualidades perceptivas del sonido (Gatehouse & Noble, 2004).
STOI (<i>Short-Time Objective Intelligibility</i>)	Métrica objetiva para estimar la inteligibilidad del habla basada en la correlación entre segmentos espectrales cortos de la señal limpia y la señal procesada. Correlaciona razonablemente con evaluaciones perceptuales en condiciones de ruido estacionario.
Teleaudiología	Atención audiológica mediante tecnologías de comunicación a distancia, que incluye evaluación, adaptación, seguimiento y educación del usuario sin presencia física
Triple filtro	Marco de tres pasos para interpretar documentación de fabricantes: (a) registrar descripciones de arquitectura como información descriptiva; (b) contrastar con evidencia independiente cuando es posible; y (c) formular afirmaciones de beneficio como hipótesis clínicas sin validación externa

Anexo D. Ficha de extracción cumplimentada

La ficha siguiente muestra un ejemplo de aplicación de la plantilla de extracción y codificación descrita en la Tabla 5 al estudio de Schulz et al. (2026), tomado como caso ilustrativo.

Tabla 9.

Ficha de extracción cumplimentada para Schulz et al. (2026).

Campo	Contenido extraído
ID interno	A01
Referencia completa	Schulz, K. M. B., Wittmann, H., Hackenberg, L., Jürgens, T., & Tchorz, J. (2026). Effects of low-latency deep-learning-based noise reduction on speech intelligibility for normal hearing, hearing-impaired listeners and cochlear implant users. <i>Frontiers in Audiology and Otology</i> , 3, 1746635. https://doi.org/10.3389/fauot.2025.1746635
Tipo de fuente	Artículo científico revisado por pares (Brief Research Report). Acceso abierto (CC BY). Publicado en <i>Frontiers in Audiology and Otology</i> , 06 de enero de 2026.
Dominio	Procesamiento de señal / AP — evaluación perceptual con usuarios
Tarea primaria (eje)	(E1) Reducción de ruido y realce del habla
Tarea secundaria	No aplica
Enfoque IA	AP (aprendizaje profundo)
Representación de entrada	Señal temporal (dominio del tiempo). Red convolucional de separación de audio en canal único (Conv-TasNet), procesamiento causal con latencia ultrabaja.
Salida del modelo	Señal de habla realzada (señal temporal limpia estimada). Sin máscara intermedia explícita; la red opera directamente sobre la forma de onda.
Punto de integración	Procesamiento offline en MATLAB sobre las señales de habla en ruido, antes de presentarlas a los participantes (vía cascos a NH y HI; vía streamer Audiolink/MED-EL o TV-streamer/Cochlear a los procesadores propios de los IC). No hay integración en hardware del audífono.

Campo	Contenido extraído
Causalidad	Causal. El modelo procesa únicamente información pasada y presente, compatible con uso en tiempo real.
Latencia reportada	Latencia ultrabaja explícita: 1 ms (correspondiente a 16 muestras de look-ahead a 16 kHz, según el análisis de la respuesta al impulso del modelo). Esta cifra está muy por debajo de los umbrales habitualmente referidos para adaptaciones auditivas abiertas (5–6 ms; Stone et al., 2008).
Restricción computacional	Modelo Conv-TasNet preentrenado tomado “off-the-shelf” por los autores; discuten la viabilidad de implementación en dispositivos de baja potencia computacional, aunque el procesamiento en este estudio se realizó en un ordenador estándar (4 veces más rápido que tiempo real). No se especifican requisitos de hardware concretos para audífonos.
Robustez	Evaluación conductual en un único tipo de ruido (olnoise/speech-shaped noise estacionario) y en SNR variable adaptativo (start +10 dB, ajustado para alcanzar 50 % de inteligibilidad). Las predicciones STOI se calcularon adicionalmente con olnoise y babble noise (cafetería). No se evalúan ruidos no vistos en entrenamiento ni distintos ajustes de audífono.
Comparador	Condición sin procesamiento (señal sin reducción de ruido) frente a condición DNN; STOI utilizada como referencia instrumental para predicción del beneficio.
Tipo de evaluación	Mixta: objetiva (STOI) y perceptual con oyentes reales (3 grupos).
Métricas / resultados primarios	SRT (umbral de recepción del habla en ruido, en dB SNR): medida primaria de inteligibilidad. STOI: métrica instrumental comparativa. Análisis de diferencias individuales y predicción del beneficio en función del SRT no procesado.
Población	51 adultos en 3 grupos: (a) oyentes con audición normal (NH, n = 20, edad 21–54 años, media 27,8), (b) personas con hipoacusia leve a moderada usuarias de audífono (HI, n = 19, edad 36–78 años, media 67,5, experiencia >2 años), (c) usuarios postlinguales de implante coclear (n inscrito = 12, n analizado = 10, edad 23–82 años; 2 participantes excluidos por problemas de conexión con el streamer).
Condiciones experimentales	OLSA (Oldenburger Satztest, frases en alemán). Para la prueba conductual: olnoise (ruido conformado al espectro del habla, estacionario) con SNR adaptativo. NH a

Campo	Contenido extraído
	65 dB(A); HI a 80 dB(A); IC con nivel ajustado a sonoridad comfortable mediante streaming. Para el cálculo de STOI: 120 frases $\times$ 33 SNR (-16 a $+16$ dB en pasos de 1 dB), procesadas en olnoise y en babble noise, totalizando 3.960 frases por condición de ruido.
Conclusión principal	El procesamiento basado en AP de baja latencia degrada ligeramente la inteligibilidad en NH (mediana SRT $-0,9$ dB), pero la mejora en HI (mediana $0,8$ dB) y de forma marcada en IC (mediana $5,7$ dB). El beneficio es mayor cuanto peor es el SRT no procesado. STOI predice razonablemente el beneficio medido en inteligibilidad, con sobreestimación de $2-3$ dB.
Limitaciones declaradas	Modelo preentrenado en inglés y aplicado a alemán (transferibilidad). Evaluación conductual en un único tipo de ruido (olnoise). Tamaño muestral reducido en grupo IC (n analizado = 10). No se evaluó escucha en condiciones ecológicas reales, sonoridad subjetiva ni esfuerzo de escucha. Los IC usaron sus procesadores propios (variabilidad de dispositivos, compresión y ancho de banda).
Riesgo de sobreinterpretación	Bajo: los autores son explícitos en los límites de generalización. El diseño es técnicamente sólido y la evaluación con los 3 grupos aporta relevancia clínica directa. La restricción a un solo ruido conductual y la ausencia de seguimiento ecológico limitan la extrapolación a uso cotidiano. El procesamiento es offline y no implementado en hardware de audífono.

Nota. SRT = umbral de recepción del habla (speech reception threshold); SNR = relación señal-ruido; NH = normal hearing; HI = hearing impairment; RNP = red neuronal profunda.