

Consensus Auditory Perceptual Evaluation of Voice (CAPE-V): revisión sistemática de los métodos utilizados para su adaptación y validación

Consensus Auditory Perceptual Evaluation of Voice (CAPE-V): Systematic Review of Methods Used for Adaptation and Validation

María-Soledad Narea-Veas¹  , Patricia Guadalupe Farías¹  , Patricia Vázquez Fernández²  

¹ Doctorado en Fonoaudiología; Universidad del Museo Social Argentino; Buenos Aires; Argentina.

² Laboratorio de Investigaciones Fonoaudiológicas; Instituto de Investigación; Universidad del Museo Social Argentino; Buenos Aires; Argentina.



Correspondencia

María-Soledad Narea-Veas.
 Email: soledad.narea@uv.cl

Citar así

Narea-Veas, María-Soledad; Farías, Patricia Guadalupe; Vázquez Fernández, Patricia. (2023). Consensus Auditory Perceptual Evaluation of Voice (CAPE-V): revisión sistemática de los métodos utilizados para su adaptación y validación. *Revista de Investigación e Innovación en Ciencias de la Salud*. 5(2), 178-204. <https://doi.org/10.46634/riics.206>

Recibido: 04/05/2023

Revisado: 09/07/2023

Aceptado: 27/09/2023

Editor

Fraidy-Alonso Alzate-Pamplona, MSc., 

Copyright

© 2023. Fundación Universitaria María Cano. La *Revista de Investigación e Innovación en Ciencias de la Salud* proporciona acceso abierto a todo su contenido bajo los términos de la licencia [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International \(CC BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Declaración de intereses

Las autoras han declarado que no hay conflicto de intereses.

Resumen

Introducción. CAPE-V es ampliamente utilizado para evaluación perceptual vocal y ha sido adaptado y validado en múltiples idiomas. A través de un análisis exhaustivo, este estudio buscó avanzar en establecer un estándar en el método utilizado para su adaptación y validación.

Objetivo. Revisar los artículos publicados entre 2002 y 2022 que han adaptado y validado CAPE-V a distintos idiomas, para evaluar exhaustivamente la adaptación, metodología y estadísticas utilizadas.

Métodos. Se realizó una revisión sistemática utilizando Scopus, Google Scholar y PubMed para identificar artículos que adaptaran y/o validaran CAPE-V entre 2002 y 2022. Se analizó el título y resumen para preseleccionar la muestra. Para evaluar el riesgo de sesgo de los estudios incluidos se analizó de forma crítica el texto completo.

Resultados. La búsqueda inicial identificó 568 artículos. Al eliminar duplicados se revisaron 559 y 23 fueron preseleccionados. 12 se incluyeron finalmente y fueron analizados, considerando adaptación, metodología y análisis estadístico. Los resultados revelan que la adaptación y validación de CAPE-V a diversos idiomas es fundamental para garantizar mediciones precisas y confiables en diferentes poblaciones.

Análisis y discusión. Existe heterogeneidad en la forma de realizar la adaptación y validación de CAPE-V. No obstante, todos los estudios tuvieron éxito en producir resultados válidos, subrayando la importancia de estos procesos para la práctica clínica.

Conclusiones. Las adaptaciones y validaciones de CAPE-V se realizaron heterogéneamente por ausencia de un protocolo estándar. Es necesario generar orientaciones para realizar estos procesos por el aporte de esta escala a la clínica, asegurando calidad y confiabilidad de los resultados.

Disponibilidad de datos

Todos los datos relevantes se encuentran en el artículo. Para mayor información, comunicarse con el autor de correspondencia.

Financiamiento

Ninguno. Esta investigación no recibió subvenciones específicas de agencias de financiamiento en los sectores público, comercial o sin fines de lucro.

Descargo de responsabilidad

El contenido de este artículo es responsabilidad exclusiva de las autoras y no representa una opinión oficial de sus instituciones ni de la *Revista de Investigación e Innovación en Ciencias de la Salud*.

Contribución de las autoras

María-Soledad Narea-Veas: Conceptualización, curación de datos, análisis formal, investigación, metodología, validación, visualización, escritura: borrador original, escritura: revisión y edición.

Patricia Guadalupe Farías: Supervisión, validación, visualización, escritura: revisión y edición.

Patricia Vázquez Fernández: Validación, visualización.

Palabras clave

Evaluación de la voz; trastornos de la voz; validación; confiabilidad; validez; adaptación transcultural; CAPE-V.

Abstract

Introduction. CAPE-V is widely used for vocal perceptual evaluation and has been adapted and validated in multiple languages. Through an exhaustive analysis, this study sought to advance in establishing a standard in the method used in its adaptation and validation.

Objective. To review articles published between 2002 and 2022 that have adapted and validated CAPE-V in different languages, to thoroughly evaluate the adaptation, methodology, and statistics used.

Methods. A systematic review was conducted using Scopus, Google Scholar and PubMed to identify articles that adapted and/or validated CAPE-V between 2002 and 2022. The title and summary were analyzed to pre-select the sample. To evaluate the risk of bias of the included studies, the full text was critically analyzed.

Results. The initial search identified 568 items. When duplicates were removed, 559 were reviewed and 23 were pre-selected. 12 were finally included and analyzed considering adaptation, methodology, and statistical analysis. The results show that the adaptation and validation of CAPE-V to different languages is essential to ensure accurate and reliable measurements in different populations.

Analysis and discussion. There is heterogeneity in how CAPE-V is adapted and validated. However, all studies were successful in producing valid results, underlining the importance of these processes for clinical practice.

Conclusions. Adaptations and validations of CAPE-V were performed heterogeneously due to the absence of a standard protocol. It is necessary to generate guidelines to perform these processes by providing this scale to the clinic, ensuring quality and reliability of results.

Keywords

Voice assesment; voice disorders; validation; reliability; validity; cross cultural adaptation; CAPE-V.

Introducción

La producción vocal es un fenómeno de carácter multidimensional que involucra mecanismos fisiológicos, biomecánicos y aerodinámicos, y es por ello que su evaluación debe considerar la medición de todas estas dimensiones para determinar el impacto de cada una en la generación del trastorno vocal. Para esto se deben elaborar historias clínicas bien acabadas, además de efectuar una batería de exámenes que abarquen evaluación auditivo-perceptual, examen anatomo-funcional de la laringe, pruebas acústicas y aerodinámicas y autopercepción del paciente respecto del trastorno vocal [1].

Respecto a la evaluación perceptual auditiva vocal por parte del fonoaudiólogo, sin duda actualmente existe conciencia acerca de la relevancia que tiene para el estudio de la voz. No obstante, ha sido complejo para los investigadores entregar

fundamentos teóricos lo suficientemente claros que expliquen qué es la calidad vocal, lo que redundaría en que existen escasas evidencias psicoacústicas para validar técnicas de evaluación. A pesar de lo anterior, se debe reconocer la importancia que este análisis tiene para llevar a cabo el diagnóstico vocal, dada su utilidad en la clínica diaria [2].

Actualmente existen varias escalas para la evaluación perceptual auditiva de la voz, entre ellas la que creó la American Speech-Language-Hearing Association (ASHA), llamada “Consensual Auditory Perceptual Evaluation of Voice (CAPE-V)”. El objetivo de la ASHA fue elaborar normas para este tipo de evaluación, basadas en la teoría, la psicoacústica, en una escala psicométrica y en la percepción de la voz, generando un instrumento que fuera de fácil aplicación y utilizable en la práctica clínica [3]. El objetivo prioritario de CAPE-V es describir la severidad de las características perceptuales de un trastorno vocal. Secundariamente pretende aportar a la comprensión anatómico-fisiológica de las alteraciones de la voz, para determinar si se requiere aplicar otras pruebas. Está concebida como una herramienta complementaria al resto de las que se utilizan en la evaluación vocal [3].

Frente a la escala CAPE-V han surgido una serie de interrogantes, entre ellas cuál es su confiabilidad intra e intersujetos, si la escala de 100 mm. y la nomenclatura utilizada en el protocolo (leve, moderada, severa) son adecuadas, con cuánta fiabilidad puede el clínico certificar la consistencia o inconsistencia de un parámetro, o si es CAPE-V efectivamente una optimización de la escala GRBAS [4]. Para dar respuesta a estos cuestionamientos se han realizado diversos estudios, que se abordarán en la presente investigación, y que han adaptado culturalmente y/o validado el instrumento con el fin de garantizar su confiabilidad, validez y aplicabilidad en el contexto específico en el que se utilizará.

Un análisis exhaustivo de todas las publicaciones que tengan por objeto la adaptación y/o validación de CAPE-V en distintos idiomas permitirá avanzar en establecer un estándar para estos procesos. Así planteado, este estudio tiene por objetivo revisar los artículos publicados entre 2002 y 2022 acerca de la adaptación y validación de CAPE-V, sistematizando la información disponible, para evaluar de forma crítica la adaptación, metodología y análisis estadístico que se utilizaron.

Métodos

Selección y descripción de la muestra

Se llevó a cabo una búsqueda sistemática de literatura relevante respecto a la adaptación y validación de CAPE-V, publicada entre 2002 y 2022, utilizando los motores de búsqueda Scopus, Google Scholar y PubMed. Las palabras claves utilizadas fueron evaluación de la voz - trastornos de la voz - validación - confiabilidad - validez - adaptación transcultural - CAPE-V. Se emplearon los operadores booleanos AND/OR para ampliar o reducir la búsqueda, según se requiriera.

Criterios de inclusión y exclusión definidos para la búsqueda de la literatura

- Publicada entre 2002 y 2022.
- Escrita en inglés o español.
- Se incluyeron artículos de revistas y tesis de grado (maestría o doctorado). Se excluyeron capítulos de libros, presentaciones en congresos, revisiones sistemáticas y actas de congresos.

- Se incluyeron solo investigaciones acerca de adaptación y validación de CAPE-V y que consideraran como escala de referencia a GRABS. Se excluyeron estudios experimentales que incluyeran a CAPE-V u otros en los que no se utilizara GRABS.

Información técnica

A través de la presente investigación se llevó a cabo una búsqueda automática y un análisis exhaustivo de todas las publicaciones que buscaran adaptar y validar la escala CAPE-V en distintos idiomas, con la finalidad de contribuir a establecer un estándar para estos procesos. Para ello se sistematizó la información disponible y se evaluó de forma crítica la adaptación, metodología y estadísticas utilizadas.

La búsqueda se llevó a cabo entre el 20 y 26 de enero de 2023. El acceso a los textos completos de los artículos se realizó a través de recursos de acceso abierto o por medio de las bases de datos del sitio institucional de la Universidad de Valparaíso de Chile, <https://bibliotecas.uv.cl/>, y estos se gestionaron utilizando el Mendeley Reference Manager. Luego de eliminar los duplicados se procedió a analizar el título y el resumen para la selección de la muestra.

No se utilizó ninguna herramienta específica para evaluar el riesgo de sesgo. Sin embargo, los textos completos de los estudios incluidos se analizaron de forma crítica considerando los siguientes criterios:

- Metodología utilizada para la adaptación transcultural, lingüística y fonética.
- Caracterización de los profesionales involucrados en la adaptación.
- Procedimientos para la validación.
- Características de las muestras vocales.
- Cantidad y experiencia de los jueces evaluadores.
- Análisis estadístico utilizado para la validación.

Por tratarse de una revisión sistemática, se realizó un análisis de los hallazgos de cada uno de los estudios, lo que se detalla en el apartado de “Resultados”.

Resultados

La búsqueda automática realizada en Scopus, Google Scholar y PubMed arrojó un total de 568 estudios, que luego de la eliminación de duplicados se redujo a 559. Al revisar manualmente y de forma crítica títulos y resúmenes y aplicar los criterios de inclusión y exclusión, se preseleccionaron 23 artículos. Se excluyeron 2 tesis, ya que se incluyeron los artículos relacionados con ellas; otra se descartó por encontrarse en un idioma distinto al inglés o español, y lo mismo ocurrió con algunos estudios que no resultaban relevantes para esta investigación. La muestra final para la revisión quedó conformada por 12 artículos, los que fueron revisados exhaustivamente en su texto completo por parte de la investigadora principal. Cada uno fue analizado considerando el proceso de adaptación, la metodología y el plan estadístico empleado para la validación. En la [Figura 1](#) se presenta el diagrama de flujo del proceso adaptado desde PRISMA 2020 [5].

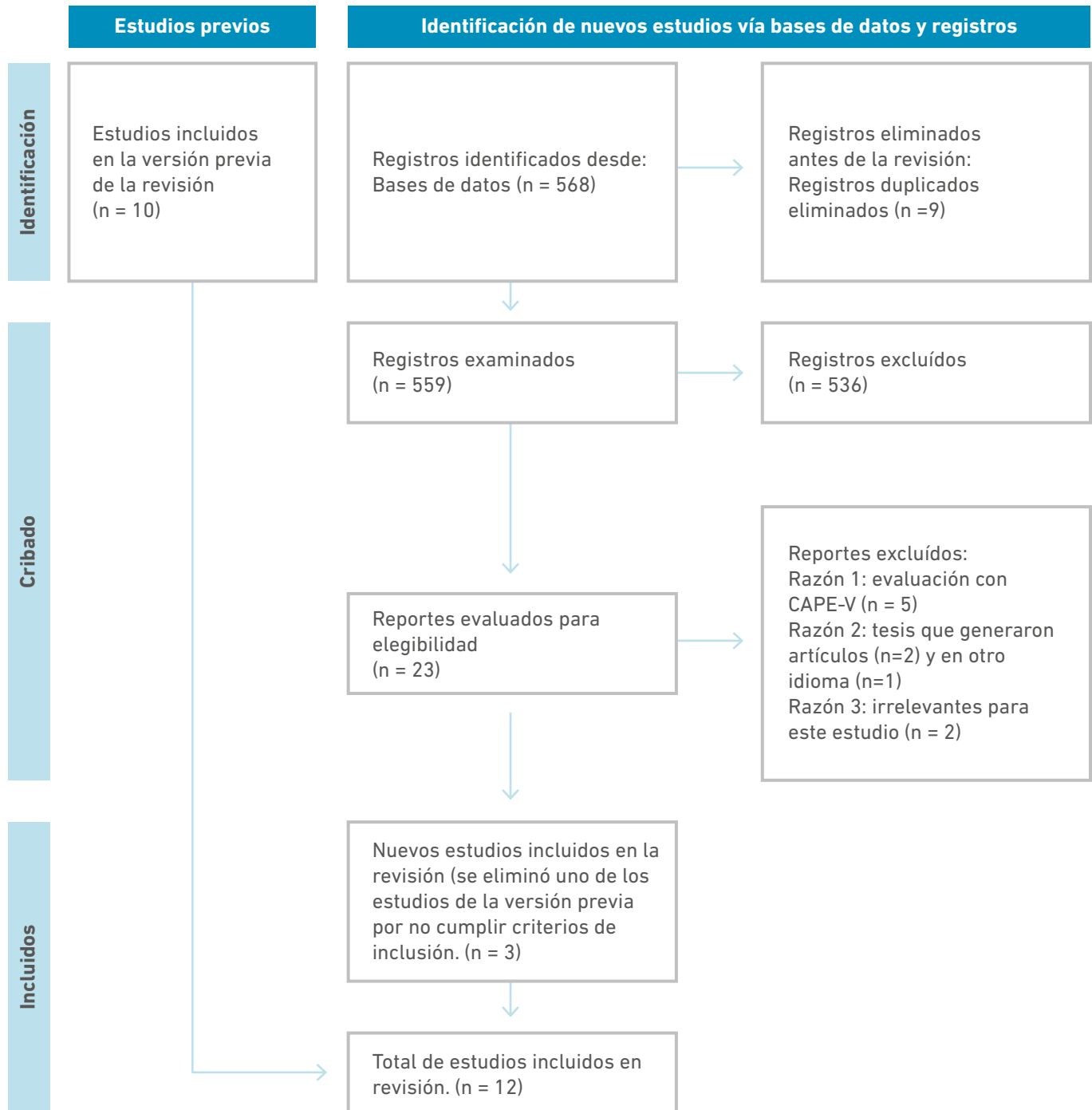


Figura 1. Proceso de búsqueda de literatura.

Nota. Fuente: adaptado de PRISMA. Flow Diagram [Internet]. Ottawa (ON): PRISMA; c2009[citado 27 de abril de 2023]. Disponible en: <http://prismastatement.org/prismastatement/flowdiagram.aspx?AspxAutoDetectCookieSupport=1>

A continuación, se presentan los resultados de los 12 artículos revisados a través de tablas. La **Tabla 1** muestra información de los estudios, ordenados por año de publicación y considerando autor(es) e idioma al que fue adaptada la escala CAPE-V. La **Tabla 2** presenta los criterios analizados en cada estudio respecto a quienes participaron en el proceso de adaptación transcultural, lingüística y fonética y qué procedimientos se utilizaron. La **Tabla 3** enseña el análisis de cada artículo en relación con el equipamiento y procedimientos utilizados para la grabación de las muestras de voz. La **Tabla 4** presenta las características de las muestras que participaron en cada estudio. La **Tabla 5** describe la cantidad y experiencia de los jueces evaluadores y los procedimientos empleados para el análisis auditivo perceptual de voces. La **Tabla 6** muestra el diseño y/o alcance y el análisis estadístico usado en cada una de las investigaciones analizadas.

Tabla 1. Información de los estudios incluidos en la revisión, ordenados por año de publicación.

Año	Autor/es	Idioma
2011	Zraïck, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., Glaze, L. E. [6]	Inglés norteamericano
2013	Mozzanica, F; Ginocchio, D; Borghi, E; Bachmann, C; Schindler, A. [7]	Italiano
2015	Núñez-Batalla, Faustino; Morato-Galán, Marta; García-López, Isabel; Ávila-Menéndez, Arántzazu. [8]	Español de España
2018	Chen, Zhen; Fang, Rui; Zhang, Yi; Ge, Pingjiang; Zhuang, Peiyun; Chou, Adriana; Jiang, Jack. [9]	Chino mandarín
2019	Özcebe, Esra; Aydinli, Fatma Esen; Tiğrak, Tuğçe Karahan; İncebay, Önal; Yilmaz, Taner. [10]	Turco
2019	de Almeida, Sancha C; Mendes, Ana P; Kempster, Gail B. [11]	Portugués europeo
2020	Ertan-Schlüter, E., Demirhan, E., Ünsal, E. M., & Tadıhan-Özkan, E. [12]	Turco
2020	Joshi, Ashwini; Baheti, Isha; Angadi, Vrushali. [13]	Hindi
2020	Gunjawate, Dhanshree R; Ravi, Rohit; Bhagavan, Srividya. [14]	Kannada
2021	Kondo, K., Mizuta, M., Kawai, Y., Sogami, T., Fujimura, S., Kojima, T., Abe, C., Tanaka, R., Shiromoto, O., Uozumi, R., Kishimoto, Y., Tateya, I., Omori, K., & Haji, T. [15]	Japonés
2022	Behlau, M., Rocha, B., Englert, M., & Madazio, G. [16]	Portugués brasileño
2022	Mossadeq, Nurhayati Mohd; Khairuddin, Khairy Anuar Mohd; Zakaria, Mohd Normani. [17]	Malayo

Nota: Artículos de investigación seleccionados en la muestra del presente estudio.

Tabla 2. Participantes y procedimientos del proceso de adaptación transcultural, fonética y lingüística del CAPE-V.

Versión	Participantes en proceso de adaptación	Procedimientos
Inglés norteamericano [6]	Versión original. Conferencia de consenso de expertos, en 2002 [3]	Consideró solo fiabilidad y validez empírica de versión original [3]
Italiano [7]	Consenso de foniatras	Consideró solo fiabilidad y validez empírica de versión adaptada al italiano publicada en otro estudio [18]. Se utilizaron tareas del protocolo original, excepto la frase que inicia con /h/ porque esta característica fonética no existe en la lengua italiana
Español de España [8]	1 logopeda	Adaptación y cambio de oraciones para permitir la valoración de los componentes de la calidad. Se consideró el contexto fonético que se pretende evaluar con cada oración original
Chino mandarín [9]	Primer autor, fonético experimental y hablante nativo de chino mandarín. Profesor universitario de fonética revisó los enunciados modificados por el autor	Adaptado utilizando los objetivos descritos por Kempster [3]. Se adaptaron 6 oraciones para fonemas de chino mandarín, considerando neutralidad semántica. Escala AV de 10 cm se aumentó agregando marcadores milimétricos y más altos cada 10 mm, transformándolo en un punto 101 para facilitar la lectura por parte de los evaluadores
Turco [10]	1 lingüista. 1 terapeuta de habla y lenguaje	Se obtuvo el permiso del departamento de derechos de autor de la ASHA. Elaboraron oraciones en turco siguiendo las reglas fonéticas descritas en la aplicación del protocolo CAPE-V
Portugués europeo [11]	1 lingüista portugués	Se realizó análisis psicométrico de la primera traducción de CAPE-V al español europeo [19] ¹ constatándose que las oraciones propuestas no consideraban todos los objetivos fonéticos de CAPE-V original. Para la retrotraducción y adaptación al portugués europeo de CAPE-V se obtuvo la autorización de la ASHA. Se elaboraron 6 nuevas oraciones, conceptualizadas y adaptadas al contexto lingüístico del portugués europeo que cumplieran con los requisitos establecidos en CAPE-V original

¹ Este estudio no fue considerado en la presente investigación por no cumplir con los criterios de inclusión.

Versión	Participantes en proceso de adaptación	Procedimientos
Turco [12]	5 terapeutas de habla y lenguaje. 1 lingüista, todos con doctorado	Para cada frase se elaboraron 5 alternativas, comparables en número de letras y sílabas con las oraciones originales de CAPE-V. Se eligieron sentencias con un número reducido de sílabas y letras. Se enviaron todas las oraciones originales del CAPE-V con transcripciones fonéticas y se solicitó a los participantes que seleccionaran 1 de las 5 traducidas al turco que mejor cumpliera el objetivo de la tarea. Se seleccionó la frase más elegida para cada tarea
Hindi [13]	1 lingüista con experiencia en hindi	Oraciones fueron elaboradas para cumplir con los requisitos de las del CAPE-V original. Se calculó el porcentaje de sonidos objetivo para cada una de las sentencias en inglés y se asimiló para las oraciones en hindi
Kannada [14]	No se menciona quienes participaron	Se obtuvo el permiso del departamento de derechos de autor de la ASHA. Adaptación no implicó una traducción del inglés al canarés. Sí fue necesaria la adaptación y el cambio de las frases para incluir todos los contextos fonéticos de la versión original de CAPE-V
Japonés [15]	Especialistas japoneses en fonética	Se obtuvo el permiso de la ASHA para desarrollar la versión japonesa del CAPE-V. Versión constaba de las mismas tareas que la versión original. Se elaboraron 6 frases específicas que contuvieran los diferentes comportamientos laríngeos. Se adaptaron a contextos japoneses de acuerdo con los mismos conceptos de la versión original
Portugués de Brasil [16]	Logopedas participantes del Grupo Reciclagem e Atualização Clínico-Científica (RACC) del Centro de Estudios de la Voz (CEV) [16]	Adaptación de CAPE-V al portugués brasileño fue realizada el año 2003 [19]. Este artículo se refiere a la validación de esa versión
Malayo [17]	4 traductores (3 logopedas y 1 lingüista), todos hablantes de inglés y malayo que cumplieran con criterios de inclusión (sin antecedentes de trastornos cognitivos, de habla o de lenguaje; y en el caso de los logopedas debían atender a lo menos un caso semanal de trastorno vocal, y debían estar familiarizados con el CAPE-V)	1 logopeda (con 21 años de experiencia) y 1 lingüista (5 años de experiencia en fonética malaya) realizaron la traducción directa. Lingüista tradujo las oraciones y verificó que las frases en malayo cumplieran con el contexto fonético de las del CAPE-V original. 2 logopedas efectuaron la retrotraducción de todo el protocolo, exceptuando las oraciones. En ambos procesos de traducción se llegó a un consenso de forma posterior a una discusión de los traductores [17]

Nota: Artículos de investigación seleccionados en la muestra del presente estudio.

Tabla 3. Descripción de equipamiento y procedimientos utilizados para la grabación de las muestras empleadas en la validación de las distintas versiones de CAPE-V.

Versión	Equipamiento y procedimientos para grabación
Inglés norteamericano [6]	<p>KayPentax Computerized Speech Lab Modelo 4500. Frecuencia de muestreo de 22 kHz.</p> <p>Micrófono de diadema (AKG Modelo C420) a una distancia de 5 cm de la boca, con ruido ambiente mínimo.</p> <p>Voces normales se grabaron igual, pero con micrófono de mano (Shure modelo SM48)</p>
Italiano [7]	<p>Señal de voz se grabó y almacenó directamente en el ordenador central utilizando el Computerized Speech Lab Modelo 4500 (Kay Elemetrics, Lincoln Park, N.J., EE. UU.) en sala con ruido ambiente <50 dB.</p> <p>Micrófono a 10 cm de la boca a 45-90°.</p> <p>Se masterizaron en DVD</p>
Español de España [8]	<p>No se especifica</p>
Chino mandarín [9]	<p>KayPentax Computerized Modelo 4500 con una frecuencia de muestreo de 22 kHz.</p> <p>Micrófono dinámico (Shure PG48) montado en un soporte de altura regulable a una distancia de 15 cm de la boca.</p> <p>Ruido ambiente mínimo. No se normalizó la intensidad de las grabaciones</p>
Turco [10]	<p>Programa Analysis of Dysphonia of Speech and Voice (CSL Model 4500, Kay Elemetrics Group, Lincoln Park, NJ, EE. UU.). Frecuencia de muestreo 25.000. Se guardaron en formato .wav.</p> <p>Micrófono de diadema Micromic C520, a una distancia de 5 cm de la boca y un ángulo de 45°.</p> <p>Sala con ruido <50 dB.</p> <p>Almacenamiento en dispositivo USB</p>
Portugués europeo [11]	<p>Grabadora digital portátil TASCAM DR-05 (Tokio, Japón), 16 bits, mono, con frecuencia de muestreo de 44.100 Hz.</p> <p>Micrófono de diadema PYLE PMEMI (China), condensador electret, omnidireccional con una respuesta en frecuencia de 20 Hz-20KHz y una sensibilidad de 44 dB±3 dB. a una distancia constante de 4 cm de la boca, en un ángulo de 45°.</p> <p>Participantes estaban sentados en una posición cómoda.</p> <p>Ruido ambiente <50 dB, medido con un sonómetro digital Rolls SLM305.</p> <p>Equipo se probó y calibró con tono puro de referencia de 500 Hz que se grababa y analizaba al principio de la grabación.</p> <p>Muestras de voz no se normalizaron para reducir la intensidad y el ruido</p>

Versión	Equipamiento y procedimientos para grabación
Turco [12]	<p>Ordenador (MacBook Pro-11,1-Mac OS X) con 16 bits de resolución y una frecuencia de muestreo de 44.000 KHz. Fueron capturadas usando el Praat (Versión 6.0.26).</p> <p>Sujetos sentados cómodamente, en una habitación silenciosa.</p> <p>Micrófono de condensador (Samson Technologies, Hicksville, Nueva York) a 4 cms de la boca en posición de 45°</p>
Hindi [13]	<p>Grabación de voces normales con Pentax Computerized Speech Laboratory y de voces disfónicas con Marantz 6000.</p> <p>Grupo voces normales: grabó las frases del CAPE-V original en inglés y habla espontánea. Ambos grupos grabaron seis frases y habla espontánea en hindi. Orden de las tareas se contrabalanceó para evitar un efecto del orden.</p> <p>Muestras grabadas en inglés e hindi se codificaron por separado para mantenerlas independientes una de la otra</p>
Kannada [14]	<p>Grabaciones en un laboratorio insonorizado con el Computerized Speech Lab Modelo 4150 (Kay Elemetrics Corp), con una frecuencia de muestreo de 50000 Hz.</p> <p>Se grabó siguiendo un procedimiento estándar [1] con un micrófono de condensador colocado a unos 10 cm de la boca en un ángulo de 45°.</p> <p>Grabaciones se almacenaron en un dispositivo de disco duro externo</p>
Japonés [15]	<p>Muestras de voz se grabaron en formato WAV con una frecuencia de muestreo de 44,1 kHz y una cuantización de 32 bits., en una sala con un nivel de ruido inferior a 50 dB.</p> <p>Micrófono de condensador (Audio-Technica AT4050, Shure SM35-XLR, AKG C747), colocado a 4 cm de la boca</p>
Portugués de Brasil [16]	<p>Ordenador MacBook Pro con el programa Audacity (versión 2.0.6) y las muestras se digitalizaron a una frecuencia de 44,1 kHz y una resolución de 16 bits.</p> <p>Micrófono de condensador AKG C420, montado en la cabeza y conectado a la interfaz de audio Focusrite iTrack solo mediante el cable AKG MPA V L.</p> <p>Participantes sentados</p>
Malayo [17]	<p>Programa de grabación Audacity, instalado en un ordenador portátil.</p> <p>Micrófono de diadema unidireccional, Sennheiser PC 8 con supresión de ruido, a una distancia boca-micrófono de 3-4 cm y un ángulo de 45° a 90°.</p> <p>Sala con un ruido ambiente inferior a 50 dB</p>

Nota: Artículos de investigación seleccionados en la muestra del presente estudio.

Tabla 4. Características de las muestras utilizadas para la validación de las diversas versiones de CAPE-V.

Versión	Muestras
Inglés norteamericano [6]	<p>Voces disfónicas: n=37 (30 mujeres y 7 hombres).</p> <p>5 evaluadores consideraron que 13 eran leves, 11 moderadas y 13 graves.</p> <p>Voces normales: n=22 (6 masculinas y 16 femeninas) de voluntarios sanos sin antecedentes de patología vocal, comprobadas perceptualmente</p>
Italiano [7]	<p>Voces disfónicas: n=80 voces (30 hombres y 50 mujeres), con 6 diagnósticos etiológicos.</p> <p>Voces normales: n=120 (48 varones y 72 mujeres), todos usuarios no profesionales de la voz y sin antecedentes de trastorno vocal</p>
Español de España [8]	<p>Voces disfónicas: n=50 (24 mujeres y 26 hombres). Se creó subgrupo de n=22 (12 mujeres y 10 hombres) sometidos a tratamiento quirúrgico, con grabación de voz pre y postoperatoria (no antes de 2 meses desde la operación).</p> <p>Voces normales: n=17 (12 mujeres, 5 hombres). Criterios de exclusión: ser menor de edad o estar o haber estado diagnosticado de alguna patología laríngea.</p> <p>Ambas muestras se aleatorizaron y se clasificaron de forma independiente</p>
Chino mandarín [9]	<p>Voces disfónicas: n=60 (29 hombres, 31 mujeres). Extraídas aleatoriamente de pacientes con queja principal de disfonía, con etiologías múltiples.</p> <p>Voces normales: n=20 (9 varones y 11 mujeres). Sujetos nunca habían sido diagnosticados de disfonía.</p> <p>Terapeuta de habla y lenguaje con 6 años de experiencia determinó gravedad general de las 80 muestras.</p> <p>10 muestras (6 disfónicas y 4 normales) se utilizaron para familiarizar a los evaluadores.</p> <p>20 muestras se repitieron para calcular la fiabilidad intrasujeto.</p> <p>70 restantes se destinaron a análisis de la fiabilidad entre evaluadores</p>
Turco [10]	<p>Grupo de estudio: n = 76. Pacientes con diversos diagnósticos. Criterios de inclusión: con buenos conocimientos de lectura, no haber sido tratados previamente por disfonía, no padecer ninguna enfermedad neurológica ni pérdida de audición, y no estar tomando ningún medicamento.</p> <p>Grupo de control: n = 54. Sin trastorno vocal. Criterios de inclusión: no debían presentar quejas de voz ni antes ni el día de la evaluación o estar resfriados, ninguna enfermedad neurológica o sistémica que pudiera afectarla voz, no haber fumado cigarrillos durante un mínimo de 5 años, tener una puntuación de 7 o inferior en el Voice Handicap Index-10, y tener buenas habilidades de lectura. 1 terapeuta de habla y lenguaje, con experiencia en trastornos de la voz, confirmó audio perceptualmente que la voz era normal</p>

Versión	Muestras
Portugués europeo [11]	<p>Voces disfónicas. N=10. Criterios de inclusión: presencia de trastorno laríngeo orgánico o funcional confirmado por laringoscopia indirecta; hablante nativo de portugués europeo; mayor de 18 años; y con capacidad de lectura y escritura adecuada para fines del estudio.</p> <p>Voces normales (control): n=10. Criterios de inclusión: ausencia de trastorno laríngeo funcional confirmado por laringoscopia indirecta; hablante nativo de portugués europeo; mayor de 18 años; y con capacidad de lectoescritura adecuada para los fines del estudio.</p> <p>Muestras no aleatorias, simples convenientes</p>
Turco [12]	<p>Voces disfónicas: n=99 (55 mujeres, 44 varones). Con diagnóstico laringológico, etiología diversa.</p> <p>Voces normales: n=83 (41 mujeres, 42 varones). No profesionales de la voz, sin antecedentes de trastorno vocal, sin infección de las vías respiratorias superiores el día de la grabación</p>
Hindi [13]	<p>Voces disfónicas: n=13, hablantes monolingües de hindi (10 hombres, 3 mujeres)</p> <p>Se les diagnosticó una patología laríngea por Un otorrinolaringólogo.</p> <p>Voces normales: n=33 adultos que hablaban hindi e inglés con fluidez; no fumadores y sin antecedentes de trastornos auditivos, del habla y del lenguaje. Se incluyeron en el estudio basándose en su autoinforme y en una evaluación perceptiva informal de su calidad vocal por parte del personal del estudio.</p> <p>Grupo normativo de participantes se reclutó en Houston y el con trastornos de la voz en la India</p>
Kannada [14]	<p>Grupo de estudio, con disfonía de diversa etiología: n = 21 (9 hombres y 12 mujeres). Criterios de inclusión: presencia de trastornos laríngeos funcionales/orgánicos confirmados mediante diagnóstico otorrinolaringológico; hablantes nativos de lengua canarés; y capacidad para leer lengua canarés.</p> <p>Grupo control, asintomático: n=44 (12 hombres y 32 mujeres). Criterios de inclusión: hablantes nativos de lengua canarés; capacidad de leer en kannada; ausencia de antecedentes de problemas de voz, enfermedades sistemáticas/neurológicas que pudieran afectar a su voz y consumo de tabaco/alcohol; y ausencia de resfriado/tos el día de la grabación. Antes de la grabación, 2 logopedas experimentados calificaron la voz de estas personas como perceptualmente normal</p>
Japonés [15]	<p>Voces disfónicas: n=76.</p> <p>Voces normales: n=97.</p> <p>Todos nativos de japonés</p>
Portugués de Brasil [16]	<p>Voces normales: n=10.</p> <p>Voces disfónicas: n=30 (10 leves, 10 moderadas y 10 graves).</p> <p>Comité de 3 logopedas, especialistas en voz, con más de 8 años de práctica clínica, seleccionó 40 voces con diferentes grados de desviación de una base de datos de 300 entre individuos disfónicos y vocalmente sanos. Se requirió el consenso de los 3 para incluir seleccionar las muestras.</p> <p>No se controló el diagnóstico ni la edad de los participantes</p>

Versión	Muestras
Malayo [17]	<p>Voces disfónicas: n=19, con diversos diagnósticos.</p> <p>Voces normales: n=19. A sujetos se les evaluó con un cuestionario de antecedentes médicos y de voz, adaptado [20].</p> <p>Inicialmente se reclutaron 47 individuos mediante un muestreo aleatorio no probabilístico. Al aplicar criterios de inclusión y exclusión por parte del autor principal (patólogo de habla y lenguaje con 12 años de experiencia) el total de participantes quedó en 38.</p> <p>Ningún participante informó dificultades cognitivas, del habla, del lenguaje o auditivas que pudieran afectar su desempeño</p>

Nota: Artículos de investigación seleccionados en la muestra del presente estudio.

Tabla 5. Descripción de la cantidad y experiencia de los jueces evaluadores y de procedimientos utilizados en el análisis auditivo perceptual de voces para la validación de las distintas versiones de CAPE-V.

Versión	Descripción de jueces	Procedimientos para análisis de voces
Inglés norteamericano [6]	<p>21 logopedas certificados por la ASHA, con experiencia en evaluación y tratamiento de personas con trastornos de la voz.</p> <p>Con más de 5 años de experiencia; un número importante de casos atendidos semanalmente; hablantes nativos de inglés; sin antecedentes de trastornos cognitivos, de habla, voz, lenguaje, audición o visión; familiaridad con las escalas CAPE-V y GRBAS; y disposición para completar las tareas de evaluación en un plazo de 72 horas</p>	<p>Se entregó CD con 74 muestras de voz (37 disfónicas, 22 normales), 11 repetidas (7 disfónicas y 4 normales) elegidas al azar para evaluar la fiabilidad intrasujeto, y 4 (2 disfónicas y 2 normales) para familiarizarse con la tarea. No se proporcionó información de ninguna de ellas.</p> <p>Evaluadores se asignaron aleatoriamente: en la primera sesión 11 juzgaron las voces (muestra de habla conversacional) con GRBAS y en la segunda con CAPE-V; 10 usaron CAPE-V en la primera y GRBAS en la segunda. Las sesiones estuvieron separadas por 48-72 horas. Escucharon las muestras de voz a campo libre con un descanso (de 5 a 10 minutos) después de la pista 36. Podían escuchar una muestra de voz más de una vez</p>

Versión	Descripción de jueces	Procedimientos para análisis de voces
Italiano [7]	<p>3 logopedas licenciados del mismo centro, especializados en evaluación y trastornos de la voz</p> <p>Con más de 5 años de experiencia.</p> <p>Ninguno participó en la grabación de la muestra de voz</p>	<p>Asistieron previamente a sesión de entrenamiento en donde evaluaron muestras de anclaje que no eran parte del estudio.</p> <p>Evaluación perceptiva de la voz se realizó a ciegas.</p> <p>Para la evaluación de la validez del CAPE-V, las puntuaciones obtenidas con las voces disfónicas se compararon con las del grupo de voces normales. Para evaluar la capacidad del CAPE-V de discriminar entre diagnósticos se compararon los 6 grupos de voces patológicas. Para validez concurrente del CAPE-V se compararon las puntuaciones obtenidas durante la primera sesión con los resultados obtenidos en la escala GRBAS.</p> <p>Se utilizaron las mismas voces para evaluar con CAPE-V y GRABS. Evaluación con GRBAS se realizó una semana después de la segunda evaluación con CAPE-V.</p> <p>Se les permitió escuchar las muestras más de una vez y, si era necesario, tomar un descanso</p>
Español de España [8]	<p>2 especialistas familiarizados con ambos métodos</p>	<p>Las voces se aleatorizaron y almacenaron en dos secuencias diferentes.</p> <p>El primer evaluador calificó con GRABS, solo escuchando la vocal sostenida. Una semana después hizo lo mismo con CAPE.V. El segundo evaluador solo calificó con CAPE-V.</p> <p>Procedimiento se realizó en un ambiente exento de ruido, escuchando desde los altavoces del ordenador o auriculares en bloques de 20, con un descanso de 10 minutos entre ellos.</p> <p>Ambos recibieron 2 patrones de escalas de voces que fueron construidos utilizando una ayuda visual espectrográfica para voces aéreas y voces ásperas.</p> <p>Podían escuchar las voces cuantas veces fuera necesario. No conocían el diagnóstico del paciente, ni tampoco que había voces normales entre las disfónicas para no influir en la fiabilidad de las evaluaciones</p>

Versión	Descripción de jueces	Procedimientos para análisis de voces
Chino mandarín [9]	<p>3 otorrinolaringólogos y 1 logopeda de 3 hospitales de China.</p> <p>Con una media de 15,25 años de experiencia clínica y realizando evaluaciones a diario</p>	<p>Evaluadores se distribuyeron aleatoriamente, 2 en el grupo A y 2 en el B. Tras la extracción de las 10 muestras de familiarización, las 90 muestras de voz se distribuyeron aleatoriamente en 2 sesiones (45 en cada una). Grupo A evaluó en la sesión 1 con GRBAS y en la 2 con CAPE-V. Grupo B calificó en la sesión 1 con CAPE-V y en la sesión 2 con GRBAS. GRBAS se evaluó con muestras de habla conversacional.</p> <p>Escucha se realizó con auriculares en una sala silenciosa. Volúmenes se ajustaron a niveles cómodos para cada oyente.</p> <p>Podían escuchar una muestra de voz varias veces. No se especificó el intervalo entre ambas sesiones, pero debían calificar con las dos sesiones secuencialmente</p>
Turco [10]	<p>2 otorrinolaringólogos y 2 terapistas del habla y lenguaje con experiencia en el campo de los trastornos de la voz</p>	<p>2 evaluadores calificaron primero con GRBAS, y 2 con CAPE-V.</p> <p>Escucharon las grabaciones a campo libre para minimizar el ruido de fondo. Utilizaron un MacBook Air. Podían ajustar el volumen a un nivel cómodo y constante y reproducir las grabaciones tantas veces como quisieran.</p> <p>Fueron ciegos a información de los participantes y a los diagnósticos del grupo de estudio.</p> <p>Se utilizaron 4 muestras de anclaje con 2 voces sanas y 2 disfónicas, que eran parte del estudio, antes de cada sesión. Se les entregaron aleatoriamente las grabaciones de los grupos de estudio y de control.</p> <p>Completaron sus evaluaciones en 2 sesiones separadas con un mínimo de 48-72 horas.</p> <p>Para el análisis de fiabilidad el 15,50% (n = 20) de las grabaciones se obtuvieron después de un mínimo de 1 semana.</p> <p>Muestras que se utilizaron para determinar la fiabilidad entre evaluadores se eligieron al azar e incluyeron muestras de los grupos de control y de estudio</p>

Versión	Descripción de jueces	Procedimientos para análisis de voces
Portugués europeo [11]	<p>14 patólogos de habla y lenguaje con más de 5 años de experiencia clínica en voz; límites auditivos bilaterales normales (medidos con audiometría); experiencia en el uso de CAPE-V y GRABS; y hablante nativo de portugués europeo</p>	<p>20 muestras de voz se presentaron una vez a cada evaluador. 6 de ellas se repitieron, mezcladas aleatoriamente con las anteriores, para determinar la fiabilidad intrasujeto.</p> <p>Durante la primera sesión de escucha los 14 evaluadores valoraron 26 muestras de voz utilizando el II EP CAPE-V. Una semana más tarde calificaron la misma secuencia de voces con GRBAS.</p> <p>Escucha se realizó en una sala silenciosa con ruido ambiente <50 dB, desde un ordenador, equipado con auriculares AKG K101. Se les permitió ajustar el volumen a un nivel de escucha cómodo, pudiendo escuchar las muestras hasta 3 veces.</p> <p>Voces se reprodujeron en 4 bloques, con un intervalo de 10 minutos entre cada uno</p>
Turco [12]	<p>2 patólogos de habla y lenguaje con doctorado, con más de 5 años de experiencia y que atendían semanalmente a pacientes con trastornos de la voz.</p> <p>Su formación académica y profesional era muy similar, y estaban familiarizados con las escalas CAPE-V y GRBAS</p>	<p>Grabaciones de voces disfónicas y normales se distribuyeron aleatoriamente y se enviaron a evaluadores.</p> <p>Se realizó sesión previa de entrenamiento con 4 muestras de anclaje de voces sanas (hombre y mujer) y disfónicas (disfonía leve y otra severa) no incluidas en el estudio. Los expertos calificaron las 178 voces restantes.</p> <p>Se tomó un descanso cada 25-30 grabaciones. Se les permitió escuchar una pista más de una vez, si era necesario.</p> <p>Para evaluar la fiabilidad intersujetos del CAPE-V los jueces calificaron todas las grabaciones y todas las tareas de escucha 2 veces en las 2 sesiones, con un intervalo de una semana.</p> <p>Lo mismo se realizó con GRBAS una semana después de la segunda sesión con CAPE-V. Existió un intervalo de 1 semana entre las 2 sesiones para evaluar la fiabilidad intrasujeto de la escala GRBAS</p>
Hindi [13]	<p>1 estudiante de posgrado y 1 logopeda con experiencia valoraron las muestras de voz normal.</p> <p>2 logopedas con más de 20 años de experiencia evaluaron las muestras de voz disfónica</p>	<p>Para evaluar la validez concurrente de la prueba se administró GRBAS a todos los participantes. Se asignaron aleatoriamente a los expertos, y las 2 muestras lingüísticas de un mismo participante no se calificaron consecutivamente</p>

Versión	Descripción de jueces	Procedimientos para análisis de voces
Kannada [14]	<p>3 logopedas con experiencia.</p> <p>Criterios de inclusión; más de 7 años de experiencia; hablante nativo de kannada; sin antecedentes de trastornos del habla, cognitivos, auditivos, visuales o lingüísticos; familiaridad con el uso de GRBAS y CAPE-V; y sensibilidad auditiva dentro de los límites normales, confirmada mediante audiometría de tonos puros</p>	<p>Evaluadores calificaron en sesiones separadas. Escucharon las muestras en un entorno de campo libre, pudiendo ajustar los niveles de volumen. No había restricciones en cuanto al número de reproducciones.</p> <p>Se evaluó con GRBAS y CAPE-V utilizando todas las tareas.</p> <p>Los evaluadores no tuvieron acceso a información de participantes.</p> <p>Hubo estímulos de anclaje, 2 para cada grupo, que no se incluyeron en el estudio. Para medir la fiabilidad intrasujeto se repitió el 15% de las muestras una semana después de la sesión inicial. y se incluyeron muestras de ambos grupos</p>
Japonés [15]	<p>5 especialistas (3 otorrinolaringólogos y 2 logopedas) con más de 5 años de experiencia en la evaluación de voces con la escala GRBAS, Nunca habían utilizado una escala analógica visual para la evaluación auditiva-perceptual</p>	<p>En una sala silenciosa, los jueces escucharon las muestras de voz con auriculares (AKG, K240MK II) en un ordenador personal con Windows a través de una aplicación de reproducción de sonido para mantener un volumen constante. Antes de la sesión de calificación cada juez practicó con 10 grabaciones de muestra.</p> <p>Cuando el rendimiento del sujeto no era uniforme en todas las tareas, se calificó su rendimiento global.</p> <p>Después de evaluar 20 voces, los jueces tomaron un descanso de 10 minutos.</p> <p>Para examinar la validez entre las escalas CAPE-V y GRBAS se realizó la calificación con alguna de ambas para todas las muestras de voz, con un intervalo mínimo de 1 semana.</p> <p>Más de 1 semana después, todas las muestras de voz fueron evaluadas con CAPE-V para el análisis de confiabilidad intrasujeto</p>

Versión	Descripción de jueces	Procedimientos para análisis de voces
<p>Portugués de Brasil [16]</p>	<p>9 patólogos de habla y lenguaje, especialistas en voz, familiarizados con GRBAS y CAPE-V, que no participaron en la selección de las muestras de voz.</p> <p>Tenían un mínimo de 10 y un máximo de 24 años de experiencia clínica</p>	<p>El análisis perceptivo auditivo fue realizado una vez con la escala GRBAS y otra con CAPE-V.</p> <p>Para aumentar la fiabilidad intra e interevaluador se presentaron voces de anclaje antes de evaluar, pudiendo acceder a ellas en cualquier momento. Cinco evaluadores calificaron primero con CAPE-V y luego con GRBAS. Cuatro evaluaron primero con GRBAS y después con CAPE-V. Fueron distribuidos aleatoriamente.</p> <p>No conocían el diagnóstico, sexo, edad y profesión</p> <p>Se programaron 2 sesiones de escucha, una para cada escala, con un intervalo mínimo de 48 horas y un máximo de 72 horas.</p> <p>Muestras se reprodujeron por altavoces y, si era necesario, se repetían.</p> <p>Calificaron con una única clasificación de la calidad vocal utilizando la vocal /a/ sostenida y el habla continua (frases del CAPE-V)</p>
<p>Malayo [17]</p>	<p>3 patólogos de habla y lenguaje, con 12, 16 y 18 años de experiencia clínica, con una carga de trabajo de al menos 1 caso de trastorno de la voz por semana.</p> <p>Sin historial de dificultades cognitivas, del habla, del lenguaje o que pudieran comprometer su capacidad para juzgar las muestras de voz</p>	<p>Se realizó un entrenamiento de familiarización con 18 tipos de muestras. Podían escucharlas las veces que fuera necesario.</p> <p>Se realizaron 2 sesiones de calificación, la primera inmediatamente después de la sesión de familiarización y la segunda una semana después. Cada sesión se realizó por separado para cada experto, en una habitación con un ruido ambiental inferior a 50 dB. En la primera sesión con CAPE-V, cada experto evaluó un conjunto aleatorio de muestras de voz presentadas en bloques de menos de 10, utilizando auriculares cerrados con tecnología de cancelación de ruido activa. Tuvieron un descanso de 10 minutos entre cada bloque.</p> <p>Podían ajustar el volumen de las muestras a un nivel cómodo y escucharlas tantas veces como fuera necesario.</p> <p>Después de una semana se llevó a cabo la segunda sesión de calificación con GRABS empleando los mismos procedimientos</p>

Nota: Artículos de investigación seleccionados en la muestra del presente estudio.

Tabla 6. Diseño o alcance declarado y análisis estadístico utilizado en los distintos estudios.

Versión	Diseño/alcance	Análisis estadístico
Inglés norteamericano [6]	No específica	<p>Confiabilidad intrasujeto CAPE-V: r de Pearson. Confiabilidad intrasujeto GRABS: r de Spearman. Confiabilidad intersujeto: coeficiente de correlación intraclase ICC.</p> <p>Grado de asociación entre los juicios CAPE-V y GRBAS: correlación múltiple</p>
Italiano [7]	Cuasiexperimental	<p>Confiabilidad intra e intersujeto de CAPE-V: ICC. Confiabilidad intra e intersujeto de los atributos C e I: coeficiente de Kappa y concordancia porcentual.</p> <p>Comparación de puntuaciones de CAPE-V entre 6 grupos de pacientes: prueba no paramétrica de Kruskal-Wallis. Comparación de valores de CAPE-V de pacientes con disfonía con los del grupo de control: t de Student.</p> <p>Para controlar el aumento del riesgo de errores de tipo I por el número de comparaciones evaluadas por las pruebas de Mann-Whitney: correcciones de Bonferroni y se estableció un nivel alfa por comparación en $p = 0,008$.</p> <p>Grado de asociación entre juicios CAPE-V y GRBAS: correlaciones de Spearman</p>
Español de España [8]	Prospectivo	<p>Fiabilidad inter e intraobservador, validez de criterio: ICC.</p> <p>Sensibilidad al cambio: t de Student</p>
Chino mandarín [9]	No específica	<p>Confiabilidad intrasujeto CAPE-V: r de Pearson. Confiabilidad intrasujeto GRABS: r de Spearman. Confiabilidad intersujeto: coeficiente de correlación intraclase ICC</p>
Turco [10]	Diseño prospectivo de casos y controles	<p>Confiabilidad intra e interevaluador del CAPE-V: ICC. Confiabilidad intraobservador: cada evaluador escuchó 20 muestras de voz (15,50%).</p> <p>Comparación de valores de CAPE-V del grupo de estudio con los del grupo de control y de diferentes grupos de trastornos vocales: prueba de Mann-Whitney U. Comparación de puntuaciones de CAPE-V de trastornos funcionales y estructurales: prueba de t de Student de muestras independientes.</p> <p>Para controlar el aumento del riesgo de errores de tipo I resultantes del gran número de comparaciones evaluadas por las pruebas de Mann-Whitney, se realizaron correcciones de Bonferroni y se estableció un nivel alfa por comparación en $P = 0,001$.</p> <p>Grado de asociación entre parámetros comparables CAPE-V y GRBAS: correlaciones de Spearman.</p> <p>Determinar si gravedad global podía predecir los grupos de estudio (estudio y normal): análisis de función discriminante y clasificación</p>
Portugués europeo [11]	Estudio transversal, observacional, descriptivo y comparativo	<p>Confiabilidad interevaluadores: ICC. Confiabilidad intraevaluadores: r de Pearson.</p> <p>Validez de constructo: t de Student. Validez concurrente: coeficiente de correlación multiserieal</p>

Versión	Diseño/alcance	Análisis estadístico
Turco [12]	No específica	<p>Confiabilidad intraevaluadores: r de Pearson. Confiabilidad interevaluadores: ICC.</p> <p>Validez de constructo: t de Student. Validez concurrente: r de Spearman.</p> <p>Cálculo de valores de corte de los grados de los parámetros de CAPE-V: análisis ROC</p>
Hindi [13]	Estudio de casos y controles	<p>Puntuaciones del CAPE-V en inglés y en hindi; parámetros comunes del CAPE-V en hindi y GRBAS en voces normales; parámetros comunes de CAPE-V en hindi y GRBAS en voces disfónicas; puntuaciones de los 2 evaluadores en voces normales y confiabilidad interevaluadores: r de Pearson</p>
Kannada [14]	Diseño prospectivo de estudio de casos y controles	<p>Confiabilidad intra e intersujeto CAPE-V: ICC.</p> <p>Resumen de valores de 6 parámetros (gravedad general, transpiración, intensidad, tono, rugosidad y tensión) para ambos grupos: media y desviación estándar para.</p> <p>Comparación de valores del grupo de control y de estudio para CAPE-V Kannada: prueba de Mann-Whitney U.</p> <p>Grado de asociación entre parámetros comparables de CAPE-V Kannada y GRBAS (grado de gravedad general, rugosidad, transpiración, tensión): correlación Spearman.</p> <p>Determinar si parámetro de gravedad general podía predecir entre los 2 grupos: análisis y clasificación de la función discriminante.</p> <p>Para controlar errores de tipo I derivados de comparaciones múltiples: FDR</p>
Japonés [15]	No específica	<p>Confiabilidad intra e intersujeto: ICC.</p> <p>Validez concurrente entre CAPE-V y GRBAS en 4 dimensiones perceptuales: r de Spearman</p>
Portugués de Brasil [16]	Estudio transversal observacional de corte transversal	<p>Confiabilidad intra e interevaluadores para GRBAS: coeficiente de Cohen y Fleiss Kappa. Confiabilidad intra e interevaluador CAPE-V: ICC.</p> <p>Distribución normal de las variables cuantitativas: prueba de Shapiro-Wilk.</p> <p>Representación de datos de los jueces: media y la mediana.</p> <p>Validez concurrente: r de Spearman</p>
Malayo [17]	No específica	<p>Determinar normalidad de datos: Shapiro Wilk</p> <p>Validez de constructo: prueba de Mann-Whitney U. Validez concurrente: correlación de Spearman. Validez de constructo: se analizaron 6 parámetros. Validez concurrente: se analizaron 4 parámetros de escala malaya CAPE-V y GRBAS.</p> <p>Tamaño del efecto: estimación del tamaño del efecto, r</p>

Nota: Artículos de investigación seleccionados en la muestra del presente estudio.

Discusión

Este artículo presenta una revisión sistemática de la literatura a través del análisis exhaustivo de todas las publicaciones que tuvieron como objetivo la adaptación y/o validación de CAPE-V en distintos idiomas, realizadas en el periodo 2002-2022, y pretende ofrecer elementos de juicio para establecer un estándar en cuanto a la realización de estos procesos. Se exploran temas como los procedimientos para efectuar la adaptación transcultural, la definición y características de la muestra, los procedimientos de grabación, el número y criterio para seleccionar a los jueces, los procedimientos para llevar a cabo la evaluación perceptual y el plan estadístico mínimo que se debería aplicar. En resumen, se busca contribuir a la normalización de los procesos de adaptación y evaluación de la voz, a fin de mejorar la calidad y la validez de los resultados obtenidos.

No obstante, todos los estudios analizados concluyen que la versión adaptada al idioma de CAPE-V es válida y confiable para ser utilizada en la evaluación perceptual de la voz en la respectiva población, los resultados de la revisión indican que existe gran heterogeneidad en la manera en que se llevaron a cabo la adaptación y la validación.

Con respecto a la adaptación de CAPE-V a los distintos idiomas, 3 de los 12 artículos analizados midieron la validez y confiabilidad a adaptaciones realizadas en otras investigaciones [6,7,16]. De los 9 restantes, 1 no señala cómo se llevó a cabo este proceso ni quienes participaron [14], y en el resto los procedimientos utilizados fueron muy heterogéneos. En general, se mantuvieron las 3 pruebas y se tuvo especial cuidado en conservar las características de las oraciones originales, respetando el contexto del idioma al que se tradujo. En algunas de las investigaciones se consideraron aspectos tales como neutralidad semántica [9], número de letras y sílabas de las oraciones [12], y porcentaje de sonidos objetivo por sentencia [13]. En el caso de la versión en chino mandarín [9], a la escala AV de 10 cm se le agregaron marcadores milimétricos y otros más altos cada 10 mm, convirtiéndolo en un punto 101 para facilitar su lectura [9]. De los 9 artículos que consideraron la adaptación, en 3 de ellos se obtuvo el permiso de la ASHA [10,11,15].

Cuando se menciona a los participantes, el número varía desde 1 a lo que es descrito como “consenso de expertos”. Con respecto a su formación, se incluyen foniatras, terapeutas de habla y lenguaje, expertos en trastornos de la voz, fonéticos y lingüistas [6-13,15-17]. Solo en un caso se menciona la intervención de traductores [17].

Un aspecto que llama la atención es que, dentro de la adaptación, solamente en 2 de los 9 estudios se consideró traducción directa y retrotraducción [11,17]. Al respecto, la bibliografía señala que es necesario hablar de “adaptación” más que de “traducción”, ya que este último procedimiento es parte del primero, que es mucho más amplio y que asegura que un determinado instrumento pueda, efectivamente, ser usado en una población distinta de la original. La adaptación es un proceso más complicado que involucra varias etapas, como asegurar la evaluación del mismo constructo en las culturas e idiomas implicados, elegir adecuadamente a los traductores, efectuar los cambios necesarios para que la prueba pueda usarse en el nuevo idioma, demostrar que la nueva versión es equivalente a la original, etc. [21].

En cuanto a los procedimientos utilizados para la grabación de las muestras de voz, en 1 de los 12 estudios no se menciona nada al respecto [8]. En 5 de los 12 se usó el KayPentax Computerized Speech Lab (versiones 4500 y 4150) [6,7,9,13,14]; en otro estudio se utilizó el Analysis of Dysphonia of Speech and Voice (CSL Model 4500 [10]; en 1 de los 12 estudios se empleó una grabadora digital portátil TASCAM DR-05 [11]; y en 3 se usó un ordenador portátil, 2 con el programa Audacity y 1 con el Praat Versión 6.0.26 [12,16,17]. Finalmente,

en la investigación restante no se indica en qué tipo de computador o programa se capturaron las voces [15]. En la versión hindi, en que la muestra se grabó en lugares diferentes, para las voces disfónicas se usó el Marantz 6000 [13].

En este mismo ámbito, la frecuencia de muestreo es informada en 8 ocasiones, pero es muy heterogénea: 22 kHz [6,9], 25 kHz [10], 44 kHz [12], 44,1 kHz [11,15,16], 50 kHz [14]. La resolución se informó solo en 4 oportunidades: 16 bits [11,12,16] y 32 bits [15].

Respecto a las características del micrófono utilizado, en algunos artículos no se indica nada o casi nada [7,8,13,14,17], en otros se entregan características básicas [6,9,10,15] y solo en 2 se describen con mayor detalle [11,16]. En cuanto a su tipo, 5 señalan haber usado de diadema [6,10,11,16,17], 2 de pedestal [6,9], 1 dinámico [9], 5 de condensador [11,12,14-16], 1 omnidireccional [11] y 1 unidireccional [17].

En cuanto a la distancia a la que fue colocado el micrófono respecto a la boca, en 2 estudios no se menciona [8,14], y en el resto va de 3-4 [17] o 4 cm [11,12,15], pasando por 5 cm [6,10], 10 cm [7,14] hasta llegar a 15 cm [9], que es un caso en que se especifica que este se encuentra montado en un soporte. Respecto al ángulo en el que se ubica, solo se consigna en 6 de los artículos: en 4 es de 45° [10-12,14] y en 2 se señala de 45° a 90° [7,17].

Se hace mención del ruido ambiente como mínimo [3,9], sala silenciosa o insonorizada [12,14] y <50 dB [7,10,11,15,17].

En contraste, la bibliografía revisada establece que para la grabación de muestras de voz o habla debiera usarse un micrófono de cabeza omnidireccional, debidamente calibrado, ubicado a una distancia de 4 a 10 cm de la boca, en un ángulo de 45° a 90°; con un preamplificador de ese micrófono y un sistema que permita la conversión, el almacenamiento y posterior análisis [1].

La grabación digital se puede llevar a cabo con una tarjeta interna de sonido de alta calidad de un computador o con un dispositivo analógico digital externo, que se puede combinar con un preamplificador de micrófono. No obstante, es más recomendable el sistema externo que el interno. Las especificaciones básicas son: frecuencia de muestreo $\geq 44,1$ kHz; resolución mínima de 16 bits, pero se recomiendan 24 bits para un mayor rango dinámico; nivel de ruido menor a 10 dB en comparación con las fonaciones más débiles; ganancia ajustable para evitar saturación; y un formato de archivo de audio sin compresión o sin pérdidas (el más adecuado es .wav) [22].

El nivel del ruido ambiente debe estar por lo menos 10 dB por debajo del nivel de las emisiones vocales más débiles. Para confirmar la calidad de la señal se le debe solicitar al sujeto que permanezca en silencio durante 5 segundos, momento en el cual se debe registrar este nivel [1].

Respecto a las muestras de participantes en los distintos estudios, en todos se dividió en voces normales y disfónicas, pero existió gran variedad en los procedimientos de selección y en el número de sujetos que las componían. En algunos casos se señalaba de forma explícita que se realizó una evaluación anatómico-funcional de la laringe para determinar la etiología y clasificar los trastornos de voz [10-14,17]; en otros, se informó que las muestras se extrajeron de bases de datos [6,16]. En 2 de los artículos se consideró para fines del estudio a subgrupos de acuerdo con su diagnóstico [7], o con intervenciones quirúrgicas realizadas [8]. En algunas de las investigaciones se determinó el grado de severidad de la disfonía [6,9,16].

Con respecto al número de sujetos en cada muestra, se observó heterogeneidad entre las investigaciones y al comparar los grupos de voces normales con los que presentaban disfonía dentro del mismo estudio. El grupo más pequeño de voces normales estuvo compuesto por 10 sujetos [11,16], lo mismo que el grupo disfónico [11]. La muestra normativa con mayor número de participantes tenía 97 [15], mientras que la de emisiones disfónicas contaba con 99 [12]. En cuanto a los criterios de inclusión y exclusión, 9 de los 12 artículos [6-8,10-14,17] los declararon en su totalidad o en parte.

Si se considera como modelo el utilizado en la versión original de CAPE-V, este supuso inicialmente la identificación de voces disfónicas que incluían una variedad de trastornos, edades y niveles de gravedad, extraídas de una base de datos de 200 ejemplos grabados de acuerdo con el protocolo de CAPE-V. Posteriormente, 2 oyentes los analizaron eliminando 14 por problemas técnicos. Los 186 restantes fueron evaluados por 5 jueces experimentados, utilizando una escala de Likert de 4 puntos (1 = normal, 2 = disfonía leve, 3 = disfonía moderada y 4 = disfonía severa). Finalmente, la muestra quedó compuesta por 37 voces disfónicas que lograron el consenso de los evaluadores (7 masculinas y 30 femeninas). Estas fueron clasificadas en 13 disfonías leves, 11 moderadas y 13 graves [6]. Respecto a su diagnóstico, correspondían a una serie de patologías que se organizaron de acuerdo con el Manual de Clasificación de Trastornos de la Voz-I [23]. La muestra de voces normales estuvo conformada por 22 sujetos voluntarios (6 hombres y 16 mujeres) sin antecedentes de patología vocal. Esta normalidad fue perceptualmente confirmada por el primer autor [6].

En cuanto a la cantidad de jueces evaluadores que participaron en el análisis auditivo perceptual de voces esta era muy diversa, siendo el mínimo 2 [8,10,12] y el máximo 21 [6]. En cuanto a su formación, consideró a logopedas o terapeutas de habla y lenguaje, expertos en trastornos de voz [6,7,9-17], otorrinolaringólogos [9,10,15] y estudiantes de posgrado [13]. En un caso se habla de especialistas, sin indicar mayores antecedentes [8].

Respecto a los criterios de inclusión para los jueces, en aquellos estudios en que se consideró el número de años de experiencia el mínimo fue de 5 años [6,7,11,12,15] y el máximo de 7 [14]. En 5 de ellos se indicó el promedio de años: 15,25 años [9], 20 años [13]; mientras en otros se informó el mínimo y máximo (10 y 24 años respectivamente) [16]; 12, 16 y 18 años respectivamente [17] como una característica de los jueces. En 4 ocasiones se consideró como criterio para ser parte del estudio el atender diaria o semanalmente a pacientes portadores de trastornos de voz [6,9,12,17]. Otra condición que se mencionó es la familiarización con CAPE-V y GRABS [6,11,12,14,16]. En este último caso llama la atención de un estudio que se señala la experiencia con GRABS, pero se indica que no existía conocimiento acerca del uso de una escala analógica visual para la evaluación auditiva-perceptual [15].

Nuevamente, si se considera el estudio que validó la versión original de CAPE-V, en él participaron 21 patólogos del lenguaje y habla certificados por la ASHA, con experiencia en evaluación y tratamiento de pacientes con trastornos de la voz. Dentro de los criterios de inclusión se definió que tuvieran más de 5 años de experiencia, atendieran semanalmente un número de pacientes con desórdenes vocales, fueran hablantes nativos de inglés, no presentaran antecedentes de déficit cognitivos, de habla, lenguaje, audición o visión, estuvieran familiarizados con CAPE-V y GRBAS y manifestaran disponibilidad para participar en un periodo de tiempo de 72 horas [6].

Acerca de los procedimientos que se emplearon en la evaluación perceptivo-acústica de las voces normales y disfónicas estos fueron muy variados. En todos los casos las muestras se aleatorizaron y la evaluación se realizó a ciegas y se efectuaron 2 sesiones para calificar los estímulos. Estas estuvieron separadas por intervalos de 48 a 72 horas [6,10,16] o una semana [7,8,11,12,14,15,17]. En un caso se informó que no se definió este periodo, pero sí que las sesiones debían ser secuenciales [9]. De los artículos revisados, 1 no hace mención del espacio de tiempo entre sesiones [13].

De los 12 estudios, 9 informaron la realización de una sesión de entrenamiento o la escucha de estímulos de anclaje de forma previa a la calificación [6,7,9,10,12,14,15-17]. Estas muestras no formaban parte de la investigación.

La escucha, en los casos en que figura en los textos analizados, se realizó en un ambiente exento de ruido, una sala silenciosa o con ruido <50dB [8,9,11,15,17], en ocasiones a campo libre [6,10,14], desde altavoces de un ordenador [8,16], o a través de auriculares [8,9,11,15,17]. Se consideró descansos establecidos cada cierto número de estímulos [6,8,11,12,15,17] o según necesidad [7], pudiendo escuchar las voces más de una vez [6,7,9,10,12,14,16,17], o hasta 3 veces [11].

Respecto al orden y estructura de cada sesión resultan ser muy heterogéneas, siendo complejo relevar aspectos similares. En general, se efectuó como mínimo una sesión en que se evaluó con CAPE-V y otra con GRABS.

A modo de consideración y para comparar, en la versión original para la validación de CAPE-V se entregó un CD con 74 muestras de voz (37 disfónicas, 22 normales), 11 repetidas (7 disfónicas y 4 normales) elegidas al azar para evaluar la fiabilidad intrasujeto, y 4 (2 disfónicas y 2 normales) para familiarizarse con la tarea. No se proporcionó información de ninguna pista. Los evaluadores se asignaron aleatoriamente, de tal manera que en la primera sesión 11 de ellos juzgaron las voces con GRBAS, utilizando la muestra de habla conversacional, y en la segunda lo hicieron con CAPE-V. Por su parte, los otros 10 usaron CAPE-V en la primera y GRABS en la segunda. Las sesiones estuvieron separadas por 48 a 72 horas. Respecto a la escucha, esta se llevó a cabo a campo libre, con un descanso de 5 a 10 minutos luego de la pista 36. Se permitió escuchar una muestra de voz más de una vez [6].

En cuanto al diseño o alcance de los distintos estudios, en 5 de ellos no se hace mención al respecto. En los otros 7 se señala lo siguiente: diseño prospectivo de estudio de casos y controles [10,14], cuasiexperimental [7], prospectivo [8], transversal, observacional, descriptivo y comparativo [11], de casos y controles [13] y transversal observacional de corte transversal [16].

Respecto a los análisis estadísticos realizados en cada investigación, existen algunos en común y otros que se diferencian. Por un lado, para calcular la confiabilidad intrasujeto de CAPE-V, se calculó la *r* de Pearson [6,9,11,12]; en tanto, para medir la confiabilidad intra o intersujeto de esta misma escala, en otros estudios se usó ICC [6-12,14-16]. En un estudio la validez intersujeto se calculó con *r* de Pearson [13]. Por otro lado, con el fin de medir la confiabilidad intrasujeto para GRABS, se recurrió a la *r* de Spearman [6,9]. Para la validez concurrente, se usó el coeficiente de correlación multiserie en 1 caso [11] y la *r* de Spearman en 4 [12,15-17]. La validez de constructo se evaluó con *t* de Student [11,12] o con la prueba de Mann-Whitney U [17]. Para la validez de criterio se usó también ICC [8]. En el caso del grado de asociación entre los juicios para CAPE-V y GRBAS, se recurrió a una correlación múltiple [6] o a *r* de Spearman [7,10,14].

Un instrumento es confiable cuando las valoraciones efectuadas con él arrojan resultados similares en distintos momentos, contextos o poblaciones, siempre que se apliquen idénticas condiciones [24]. En la práctica, la confiabilidad se enlaza con otro concepto, que es la validez, y que se relaciona con la medida en que una prueba posee propiedades que le permiten medir de forma consistente y precisa lo que pretende evaluar. En estudios cuyo objetivo es determinar si un instrumento, en este caso CAPE-V, es válido y confiable, debería considerarse al menos la medición de la fiabilidad intra e intersujeto, y la validez de criterio, de constructo y concurrente.

En resumen, esta revisión sistemática demostró que la adaptación y validación de CAPE-V a diversos idiomas consiste en una tarea crítica que busca garantizar la calidad en la medición de las variables de interés en diferentes poblaciones. Se encontró que todos los estudios que se incluyeron en esta investigación fueron exitosos en sus resultados, lo que permite su utilización en las respectivas poblaciones. En general, estos hallazgos enfatizan la importancia de estos procesos para asegurar mediciones precisas y confiables en diferentes contextos. No obstante, es del todo necesario señalar que se encontraron diferencias importantes en la forma en que se implementó la adaptación y validación, que hacen necesario avanzar en establecer un protocolo mínimo que permita efectivamente comparar resultados.

Conclusiones

En este estudio se evidencia que las adaptaciones y validaciones de CAPE-V han sido realizadas de manera heterogénea, debido a la falta de un protocolo estándar para llevar a cabo estos procesos. Ante el aumento de versiones, parece necesario establecer orientaciones específicas para la adaptación cultural, lingüística y fonética de instrumentos como CAPE-V. Establecer criterios de inclusión para la selección de las muestras, así como procedimientos de evaluación perceptual de las voces y medidas mínimas, permitirá realizar comparaciones o generar conclusiones más fidedignas entre las distintas investigaciones. Es fundamental sentar bases sólidas que permitan la adaptación y validación de CAPE-V en diferentes contextos culturales, debido a su gran valor en la práctica clínica. Este enfoque asegurará la calidad y confiabilidad de los resultados y permitirá el uso efectivo de esta herramienta en diversas comunidades.

Referencias

1. Patel RR, Awan SN, Barkmeier-Kraemer J, Courey M, Deliyiski D, Eadie T, et al. Recommended protocols for instrumental assessment of voice: American Speech-Language-hearing association expert panel to develop a protocol for instrumental assessment of vocal function. *Am J Speech Lang Pathol* [Internet]. 2018;27(3):887-905. doi: http://dx.doi.org/10.1044/2018_AJSLP-17-0009
2. Kreiman J, Gerratt BR. Perceptual assessment of voice quality: Past, present, and future. *Perspect Voice Voice Disord* [Internet]. 2010;20(2):62-7. doi: <http://dx.doi.org/10.1044/vvd20.2.62>
3. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol* [Internet]. 2009;18(2):124-32. doi: [http://dx.doi.org/10.1044/1058-0360\(2008/08-0017\)](http://dx.doi.org/10.1044/1058-0360(2008/08-0017))
4. Kempster G. CAPE-V: Development and future direction. *Perspect Voice Voice Disord* [Internet]. 2007;17(2):11-3. doi: <http://dx.doi.org/10.1044/vvd17.2.11>

5. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Rev Esp Cardiol* [Internet]. 2021;74(9):790-9. doi: <http://dx.doi.org/10.1016/j.recesp.2021.06.016>
6. Zraick RI, Kempster GB, Connor NP, Thibeault S, Klaben BK, Bursac Z, et al. Establishing validity of the Consensus Auditory-Perceptual Evaluation of voice (CAPE-V). *Am J Speech Lang Pathol* [Internet]. 2011;20(1):14-22. doi: [http://dx.doi.org/10.1044/1058-0360\(2010/09-0105\)](http://dx.doi.org/10.1044/1058-0360(2010/09-0105))
7. Mozzanica F, Ginocchio D, Borghi E, Bachmann C, Schindler A. Reliability and validity of the Italian version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *Folia Phoniatr Logop* [Internet]. 2014;65(5):257-65. doi: <http://dx.doi.org/10.1159/000356479>
8. Núñez-Batalla F, Morato-Galán M, García-López I, Ávila-Menéndez A. Adaptación fonética y validación del método de valoración perceptual de la voz CAPE-V al español. *Acta Otorrinolaringol Esp* [Internet]. 2015;66(5):249-57. doi: <http://dx.doi.org/10.1016/j.otorri.2014.07.007>
9. Chen Z, Fang R, Zhang Y, Ge P, Zhuang P, Chou A, et al. The Mandarin version of the Consensus Auditory-Perceptual Evaluation of voice (CAPE-V) and its reliability. *J Speech Lang Hear Res* [Internet]. 2018;61(10):2451-7. doi: http://dx.doi.org/10.1044/2018_JSLHR-S-17-0386
10. Özcebe E, Aydinli FE, Tiğrak TK, İncebay Ö, Yılmaz T. Reliability and validity of the Turkish version of the Consensus Auditory-Perceptual Evaluation of voice (CAPE-V). *J Voice* [Internet]. 2019;33(3):382.e1-382.e10. doi: <http://dx.doi.org/10.1016/j.jvoice.2017.11.013>
11. de Almeida SC, Mendes AP, Kempster GB. The Consensus Auditory-Perceptual Evaluation of voice (CAPE-V) psychometric characteristics: II European Portuguese version (II EP CAPE-V). *J Voice* [Internet]. 2019;33(4):582.e5-582.e13. doi: <http://dx.doi.org/10.1016/j.jvoice.2018.02.013>
12. Ertan-Schlüter E, Demirhan E, Ünsal EM, Tadihan-Özkan E. The Turkish version of the Consensus Auditory-Perceptual Evaluation of voice (CAPE-V): A reliability and validity study. *J Voice* [Internet]. 2020;34(6):965.e13-965.e22. doi: <http://dx.doi.org/10.1016/j.jvoice.2019.05.014>
13. Joshi A, Baheti I, Angadi V. Cultural and linguistic adaptation of the Consensus Auditory-Perceptual Evaluation of voice (CAPE-V) into hindi. *J Speech Lang Hear Res* [Internet]. 2020;63(12):3974-81. doi: http://dx.doi.org/10.1044/2020_JSLHR-20-00348
14. Gunjawate DR, Ravi R, Bhagavan S. Reliability and validity of the Kannada version of the Consensus Auditory-Perceptual Evaluation of voice. *J Speech Lang Hear Res* [Internet]. 2020;63(2):385-92. doi: http://dx.doi.org/10.1044/2019_JSLHR-19-00020
15. Kondo K, Mizuta M, Kawai Y, Sogami T, Fujimura S, Kojima T, et al. Development and validation of the Japanese version of the Consensus Auditory-Perceptual Evaluation of voice. *J Speech Lang Hear Res* [Internet]. 2021;64(12):4754-61. doi: http://dx.doi.org/10.1044/2021_JSLHR-21-00269

16. Behlau M, Rocha B, Englert M, Madazio G. Validation of the Brazilian Portuguese CAPE-V instrument-Br CAPE-V for auditory-perceptual analysis. *J Voice* [Internet]. 2022;36(4):586.e15-586.e20. doi: <http://dx.doi.org/10.1016/j.jvoice.2020.07.007>
17. Mohd Mossadeq N, Mohd Khairuddin KA, Zakaria MN. Cross-cultural adaptation of the consensus auditory-perceptual evaluation of voice (CAPE-V) into Malay: A validity study. *J Voice* [Internet]. 2022. doi: <http://dx.doi.org/10.1016/j.jvoice.2022.05.018>
18. Jesús LM, Barney A, Santos R, Caetano J, Couto JJ. Universidade de Aveiro's voice evaluation protocol. In: INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association; 2009 Sep 6-10; Brighton: International Speech Communication Association (ISCA); 2009. p. 971-4. doi: <https://doi.org/10.21437/interspeech.2009-289>
19. Behlau M, Oliveira G. Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), ASHA 2003. *Rev Soc Bras Fonoaudiol.* 2003;9. Available from: https://www.researchgate.net/publication/285981950_Consensus_Auditory-Perceptual_Evaluation_of_Voice_CAPE-V
20. Lundy DS, Casiano RR, Sullivan PA, Roy S, Xue JW, Evans J. Incidence of abnormal laryngeal findings in asymptomatic singing students. *Otolaryngology-Head and Neck Surgery* [Internet]. 1999;121(1):69-77. doi: [http://dx.doi.org/10.1016/s0194-5998\(99\)70128-2](http://dx.doi.org/10.1016/s0194-5998(99)70128-2)
21. Cardoso Ribeiro C, Gómez-Conesa A, Hidalgo Montesinos MD. Metodología para la adaptación de instrumentos de evaluación. *Fisioter (Madr, Ed, Impresa)* [Internet]. 2010;32(6):264-70. doi: <http://dx.doi.org/10.1016/j.ft.2010.05.001>
22. Švec JG, Granqvist S. Guidelines for Selecting Microphones for Human Voice Production Research. *Am J Speech Lang Pathol* [Internet]. 2010;19(4):356-68. doi: [https://doi.org/10.1044/1058-0360\(2010/09-0091\)](https://doi.org/10.1044/1058-0360(2010/09-0091))
23. Verdolini K, Rosen CA, Branski RC. Classification manual for voice disorders-I. Mahwah, NJ, Estados Unidos de América: Lawrence Erlbaum Associates; 2005. 296 p. doi: <https://doi.org/10.4324/9781410617293>
24. Manterola C, Grande L, Otzen T, García N, Salazar P, Quiroz G. Confiabilidad, precisión o reproducibilidad de las mediciones. Métodos de valoración, utilidad y aplicaciones en la práctica clínica. *Rev Chilena Infectol* [Internet]. 2018;35(6):680-8. doi: <http://dx.doi.org/10.4067/s0716-10182018000600680>