

AUDIODESCRIPCIÓN AUTOMATIZADA EN LA ENSEÑANZA DEL ESPAÑOL COMO LENGUA EXTRANJERA: DESAFÍOS Y PERSPECTIVAS

Lídia Scarabele-Elis ¹

Gleiton Malta ²

RESUMEN

A partir de 2020, se intensificó la popularización de las herramientas de IA, entendidas como tecnologías capaces de simular habilidades humanas. En este contexto, se ampliaron las discusiones sobre sus impactos en diferentes áreas, especialmente en la educación. Este artículo tiene como objetivo analizar audiodescripciones automatizadas producidas por herramientas de IA a partir de imágenes del libro didáctico de español *Gente Única*, considerando criterios de descripción de personajes, ambientación, contextualización y adecuación pedagógica. Se analizaron audiodescripciones generadas por *ChatGPT*, *Zapia*, *Be My Eyes*, *Lookout* y *Copilot*, a partir de Motta (2016) y Naves *et al.* (2016). Este es un estudio cualitativo, de carácter exploratorio, desarrollado mediante un estudio de caso (Chizzotti, 2006). Los resultados evidencian limitaciones de las IAs en cuanto a la sensibilidad pedagógica, cultural y lingüística necesaria para la audiodescripción en materiales didácticos, reforzando la importancia del profesional audiodescriptor para garantizar una accesibilidad adecuada en el contexto educativo.

PALABRAS CLAVE

Audiodescripción; Inteligencia artificial; Español como lengua extranjera; Accesibilidad.

RESUMO

A partir de 2020, intensificou-se a popularização das ferramentas IA, entendidas como tecnologias capazes de simular habilidades humanas. Nesse contexto, ampliaram-se as discussões sobre seus impactos em diferentes áreas, especialmente na educação. Este artigo objetiva analisar audiodescrições automatizadas produzidas por ferramentas de IA a partir de imagens do livro didático de espanhol *Gente Única*, considerando critérios de descrição de personagens, ambientação, contextualização e adequação pedagógica. Foram analisadas audiodescrições geradas pelas plataformas *ChatGPT*, *Zapia*, *Be My Eyes*, *Lookout* e *Copilot*, com base em Motta (2016) e Naves *et al.* (2016). Trata-se de um estudo qualitativo, de caráter exploratório, desenvolvido por meio de estudo de caso (Chizzotti, 2006). Os resultados evidenciam limitações das IAs quanto à sensibilidade pedagógica, cultural e linguística necessária à audiodescrição em materiais didáticos, reforçando a importância do profissional audiodescriptor para garantir acessibilidade adequada no contexto educacional.

¹ Mestra em Estudos da Tradução pela Universidade de Brasília (UnB). Audiodescritora na Scarabele audiodescrição. E-mail: scarabeleaudiodescricao@gmail.com. ORCID: <https://orcid.org/0009-0006-2107-8386>.

² Doutor em Estudos Linguísticos com ênfase em Estudos da Tradução pela Universidade Federal de Minas Gerais (UFMG). Professor permanente do Programa de Pós-graduação em Língua e Cultura (PPGLinC) do Instituto de Letras da Universidade Federal da Bahia (UFBA). E-mail: gleitonmalta@ufba.br. ORCID: <https://orcid.org/0000-0003-0359-4046>.

PALAVRAS-CHAVE

Audiodescrição; Inteligência artificial; Espanhol como língua estrangeira; Acessibilidade.

INTRODUCCIÓN³

En las últimas décadas, la inteligencia artificial (IA) ha ampliado su estancia en diferentes esferas sociales, incluyendo el ámbito educativo y el desarrollo de tecnologías de apoyo orientadas a la accesibilidad. Muchas herramientas integradas con IA han sido utilizadas en procesos relacionados con la comunicación, la movilidad y el acceso a la información de personas con discapacidad visual, demostrando potencial para automatizar recursos de accesibilidad, como la audiodescripción (Silva *et al.*, 2024).

En el contexto educativo, el creciente uso de estas herramientas debe ser evaluado críticamente, en especial en la enseñanza del español como lengua extranjera (ELE). Esto se debe a que los manuales didácticos contemporáneos han incorporado recursos fotográficos, ilustraciones, gráficos y viñetas con el propósito de promover la construcción de la competencia comunicativa (Chang, 2019). A parte de eso, un 25% de los estudios defendidos en Brasil en nivel de máster y doctorado, cuyo el tema es la audiodescripción, se han realizado primero en las Letras (el 35,6%) y en segundo lugar en las Ciencias sociales, principalmente en el área de la Educación (el 24,2%) (Malta, 2026). En este sentido, las herramientas de IA pueden contribuir a la generación de audiodescripciones que permitan hacer accesible dicho contenido visual. Sin embargo, la ausencia de audiodescripción o una audiodescripción inadecuada puede reproducir la exclusión del aprendiz al dificultar el acceso a las informaciones visuales, aspecto que puede derivarse del uso indiscriminado o poco criterioso de recursos de IA. Así que, los recursos multimodales presentes en los materiales didácticos, utilizados como formas de contextualización lingüística y cultural, necesitan ser audiodescritos cuidadosamente para

³ Declaramos que se utilizaron herramientas de Inteligencia Artificial Generativa en este trabajo en las siguientes condiciones: *ChatGPT (OpenAI)* fue utilizada para resumir secciones de la disertación de maestría. Entrada: Resuma este fragmento, preservando los principales argumentos y las ideas centrales. *PerplexityAI* fue utilizada para resumir para la revisión textual. Entrada: Adapte este texto al lenguaje propio de un artículo científico, observando las normas de la escritura académica y manteniendo el sentido original. *Perplexity* fue utilizada para resumir. Todo el contenido generado con el apoyo de estas herramientas fue sometido a revisión crítica por parte de los autores, quienes asumen plena responsabilidad por la originalidad, veracidad e integridad del texto final, de conformidad con la Política de Integridad en la Actividad Científica del CNPq (Ordenanza n.º 2.664/2026). Por fin, todo el trabajo pasó por una revisión especializada por profesional habilitado en normas técnicas académicas y revisión textual, hablante de español como segunda lengua.

garantizar la accesibilidad, especialmente desde la perspectiva de una intervención humana especializada (Scarabele-Elis; 2022; Malta: Scarabele-Elis; 2023; Scarabele-Elis; 2025).

Sin embargo, los estudiantes con discapacidad visual aún enfrentan dificultades para acceder a manuales didácticos accesibles, debido a la ausencia de audiodescripciones adecuadas de los contenidos visuales (Cerchiari, 2011; Medrado, 2014). La audiodescripción, entendida como un recurso “comunicacional que amplía la comprensión de las personas con discapacidad visual en todo tipo de eventos, ya sean académicos, científicos, sociales o religiosos, mediante información sonora⁴” (Motta, 2016, p. 15), constituye una importante herramienta de accesibilidad e inclusión en el contexto educativo.

Ante la creciente presencia de la inteligencia artificial en la vida cotidiana y su potencial para automatizar o auxiliar los procesos de accesibilidad, resulta pertinente investigar de qué manera estas herramientas pueden contribuir a la enseñanza y al aprendizaje de lenguas por parte de estudiantes con discapacidad visual. Además, considerando que muchos profesionales de la educación no poseen un dominio técnico adecuado sobre la interacción con herramientas generativas, conforme demuestra estudio sistemático de Galasso (2026), se hace necesario analizar el funcionamiento de estas tecnologías en contextos más próximos al uso cotidiano escolar.

En este sentido, el presente artículo, resultado de la investigación A audiodescrição realizada por ferramentas de inteligência artificial: um estudo a partir de imagens de um livro didático de espanhol como língua estrangeira (2025), llevada a cabo en el ámbito de la Maestría en Estudios de Traducción de la Universidade de Brasília, Brasil, y del grupo de investigación Mapeamentos em Tradução (MapTrad)⁵, tiene como objetivo analizar audiodescripciones automatizadas producidas por herramientas de inteligencia artificial a partir de una imagen del manual didáctico de español Gente Única, teniendo en cuenta criterios relacionados con la descripción de personajes, la ambientación, la contextualización y la adecuación pedagógica (Motta, 2016). Para ello, se analizan audiodescripciones generadas por las herramientas ChatGPT, Zapia, Be My Eyes, Lookout

⁴ Del portugués brasileño: “comunicacional que amplia o entendimento das pessoas com deficiência visual em todos os tipos de eventos, sejam eles acadêmicos, ciências, sociais ou religiosos, por meio de informação sonora.

⁵ Disponible en: <https://ppglinc.ufba.br/pt-br/maptrad-mapeamentos-em-traducao>. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) <http://dgp.cnpq.br/dgp/espelhogrupo/4862273218564576>

y Copilot, con base en las directrices propuestas por Motta (2016) y Naves et al. (2016), inteligencias artificiales que han sido utilizadas o que se proponen como recursos de accesibilidad.

Esta investigación es relevante porque resulta necesario evaluar cómo los recursos tecnológicos pueden ser aprovechados para garantizar los derechos de las personas que dependen de la accesibilidad lingüística, como el acceso a la educación. Al problematizar este tema, se obtendrán evidencias sobre el uso o la ausencia de estas tecnologías en la promoción efectiva de dichos derechos. Según el Ministerio de Salud, en referencia a la Organización Mundial de la Salud (OMS), existen aproximadamente 285 millones de personas con discapacidad visual en el mundo (2023). En Brasil, el *Instituto Brasileiro de Geografia e Estatística* (IBGE, 2022) registra 8,6 millones de personas con algún tipo de discapacidad (el 8,9 % de la población de dos años o más); de este total, el 3,1 % presenta discapacidad visual. Estas cifras exigen que el poder público implemente políticas públicas eficaces. En el ámbito educativo, el Estado tiene el deber de garantizar la escolarización de estas personas, considerando sus especificidades para asegurar el acceso y la permanencia educativa. Además, según Malta (2026), se nota un aumento creciente de investigaciones que involucran la audiodescripción en nivel de máster y doctorado en Brasil. Los estudios parten de 1 trabajo publicado en 2009, llegando a 18 en 2020 y 15 en 2022, sumando un total de 132 trabajos defendidos en 65 universidades brasileñas en el período de 2009 a 2022 abarcando diversas áreas del conocimiento.

1 El lugar de la audiodescripción en la enseñanza del español como LE

La audiodescripción (AD) constituye un campo interdisciplinario vinculado a los Estudios de Traducción, articulándose también con las áreas de la accesibilidad, la inclusión y la educación lingüística (Hurtado Albir, 2017). Insertada en el ámbito de la traducción intersemiótica, conforme a la propuesta de Jakobson (1995), la AD opera en la mediación entre signos verbales y no verbales al transformar informaciones visuales en lenguaje verbal accesible.

En el contexto educativo, la AD es un recurso imprescindible para promover la accesibilidad de estudiantes con discapacidad visual, especialmente en prácticas pedagógicas que utilizan recursos visuales, como ocurre con los materiales didácticos que han incorporado imágenes como mecanismo de contextualización. Según Motta (2016) y Motta y Romeu Filho (2010), la AD consiste en una herramienta de accesibilidad

comunicacional capaz de ampliar la comprensión de las personas con discapacidad visual mediante la traducción del contenido visual en información sonora.

Es decir, la AD transforma el lenguaje visual en verbal, garantizando “mayores posibilidades de acceso a la cultura y a la información, contribuyendo a la inclusión cultural, social y escolar⁶” (Motta, 2016, p. 15). En la misma línea de análisis, Naves *et al.* (2016, no paginado) especifican y caracterizan los elementos que constituyen la AD al afirmar que esta es “una locución adicional guionizada que describe las acciones, el lenguaje corporal, los estados emocionales, la ambientación, el vestuario y la caracterización de los personajes”⁷.

En la enseñanza del español como lengua extranjera (ELE), la relevancia de la audiodescripción se vuelve aún más evidente ante la creciente presencia de imágenes en los manuales didácticos contemporáneos (Benítez, 2009), cuyo objetivo es “convertirse en un gran instrumento que nos ofrece enormes posibilidades en la enseñanza de una lengua” (Benítez, 2009, p. 1). En este ámbito, fotografías, ilustraciones, gráficos y demás elementos visuales contribuyen a la contextualización lingüística y cultural, favoreciendo el desarrollo de la competencia comunicativa de los estudiantes (Chang, 2019). Por esta razón, se ve razonado que la falta de accesibilidad visual compromete la comprensión de los contenidos y limitan la participación de estudiantes con discapacidad visual en las actividades pedagógicas, en contravención a derechos básicos (Brasil, 2015).

Es necesario resaltar que el desarrollo de la competencia comunicativa constituye el fundamento de la enseñanza de idiomas (Cantero, 2019). Se entiende por competencia comunicativa la movilización integrada de saberes, habilidades y estrategias para negociar significados en contextos concretos y situados (Almeida Filho; Franco, 2009; Cantero, 2019). Así, todo material utilizado en el aula, especialmente los recursos visuales, sirve para representar diferentes escenarios comunicativos. Por ello, la verbalización y la descripción de estos elementos también son esenciales para los estudiantes con discapacidad visual. En palabras de Chang (2019, p. 23, el destaque es nuestro):

*Las imágenes, utilizadas como medio y apoyo para **desarrollar y consolidar la capacidad lingüística**, pueden contribuir al almacenamiento*

⁶ Del portugués brasileño: “possibilidades maiores de acesso à cultura e à informação, contribuindo para a inclusão cultural, social e escolar”.

⁷ Del portugués brasileño: “uma locução adicional roteirizada que descreve as ações, a linguagem corporal, os estados emocionais, a ambientação, os figurinos e a caracterização dos personagens”.

*en la mente como imagen mental, facilitando la reactivación lingüística en la vida real; esto significa que el sujeto utiliza, aplica o incluso re-presenta determinadas expresiones adquiridas por analogía, a través de las imágenes mentales que ha asimilado.*⁸

Es decir, la imagen posee un sentido inscrito en ella; no aparece como un elemento meramente embellecedor en el manual didáctico, sino como un soporte adicional para el desarrollo de habilidades lingüísticas y culturales, ya sea mediante un elemento que estimule la comunicación, la comprensión o incluso como mecanismo para memorizar la información. De este modo, ante la presencia de un estudiante que requiere audiodescripción, se hace evidente la necesidad de que este recurso le sea proporcionado en la enseñanza del español con el fin de garantizar la accesibilidad.

En este contexto, además de promover la inclusión y garantizar el acceso a la información, la audiodescripción se revela como un recurso importante y necesario para la mediación pedagógica en el proceso de enseñanza y aprendizaje de lenguas, dado que posibilita que los estudiantes con discapacidad visual accedan, en condiciones más equitativas, a los mismos contenidos ofrecidos a los demás alumnos (Sousa, 2017). Tal potencial deriva de la traducción de elementos visuales al lenguaje verbal, procedimiento que intensifica la construcción de sentidos, la contextualización cultural y la ampliación de las posibilidades de aprendizaje en contextos educativos marcados por la multimodalidad.

Como argumenta Motta (2016), la lectura de imágenes constituye un elemento fundamental en la contemporaneidad, especialmente en entornos educativos permeados por múltiples lenguajes y recursos semióticos. La autora (2016, p. 56) sostiene que:

*Es necesario que las imágenes y otros elementos no textuales sean descritos mediante texto, ya que una persona con discapacidad visual no puede interpretar las imágenes de la misma manera que una persona vidente. La transformación de las imágenes en palabras, en este caso, será esencial para ampliar la comprensión y promover una interacción más completa con el material didáctico.*⁹

En esta perspectiva, la AD potencia los procesos de interacción, participación y

⁸ Del portugués de Portugal: “As imagens, utilizadas como meio e auxílio para desenvolver e consolidar a capacidade linguística, podem contribuir para o armazenamento na mente, como imagem mental, facilitando a reativação linguística na vida real, isso quer dizer que o sujeito utiliza, aplica ou até re-presenta certas expressões adquiridas por analogia, através das imagens mentais que assimilou”.

⁹ Del portugués brasileño: “É preciso que as imagens e outros elementos não textuais sejam descritos em texto, já que uma pessoa com deficiência visual não pode fazer a leitura das imagens da mesma forma que uma pessoa que enxerga. A transformação das imagens em palavras, neste caso, será essencial para ampliar o entendimento e promover uma interação mais completa com o material didático”.

construcción del conocimiento en la enseñanza de lenguas. Cabe destacar que la AD ya se encuentra respaldada jurídicamente en el ordenamiento brasileño. La *Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)* (Brasil, 2015) garantiza el derecho a la accesibilidad y a la promoción de tecnologías apoyo, reconociendo la necesidad de eliminar barreras comunicacionales y pedagógicas. Más recientemente, la Ley n.º 15.249/2025 determina que los servicios de educación especializada garanticen “sistemas de comunicación aumentativa y alternativa de baja tecnología para la atención educativa especializada de estudiantes con necesidades complejas de comunicación”¹⁰ (Brasil, 2025). Aunque la normativa no menciona explícitamente la audiodescripción, se entiende que de ella deriva la ampliación del derecho a recursos mínimos de accesibilidad comunicacional, entre los cuales la AD se inserta como un instrumento fundamental.

En lo que respecta a la enseñanza del español en el contexto brasileño, la *Base Nacional Comum Curricular* (2018) sitúa la lengua española como un componente curricular de oferta preferente, lo que implica su implementación de manera heterogénea entre los Estados, las redes de enseñanza y las instituciones privadas. Aun así, toda institución que ofrezca la asignatura debe garantizar la adecuación accesible de los recursos pedagógicos utilizados, no mediante la simplificación o supresión de contenidos, sino a través de la efectiva garantía de acceso a las informaciones visuales presentes en los recursos didácticos. En este sentido, la AD constituye un recurso indispensable para la democratización del acceso al currículo. Tal comprensión ya se encuentra respaldada por la Nota Técnica n.º 21/2012, documento mediante el cual el Ministerio de Educación estableció orientaciones dirigidas a la AD, presentando específicamente los requisitos básicos para que el audiodescriptor atienda en la elaboración de su guion. Se trata de una nota orientada a las editoriales que ponen a disposición recurso didáctico digital para los públicos que dependen de la audiodescripción, específicamente en el contexto de enseñanza pública.

Ante este escenario, se destaca que el currículo debe ser concebido para todos los estudiantes, razón por la cual la AD tiene un papel central en la equiparación de los derechos educativos, al transformar contenidos imagéticos y textuales en lenguaje verbal accesible para estudiantes con discapacidad visual (Mianes, 2023). Tal relación entre texto e imagen resulta sumamente relevante en la enseñanza de lenguas, ya que los materiales

¹⁰ Del portugués brasileño: “sistemas de comunicação aumentativa e alternativa de baixa tecnologia para o atendimento educacional especializado de estudantes com necessidades complexas de comunicação”

didácticos tienden a articular elementos verbales y visuales en la construcción de sentidos, como ya reflexionado por Motta (2016). Para Lúcia Santaella (2009), textos e imágenes constituyen elementos complementarios, responsables de agregar múltiples dimensiones semióticas al proceso de aprendizaje.

Desde esta perspectiva, se comprende que la AD debe ocupar una posición estratégica en la promoción de la accesibilidad y en la garantía del disfrute de derechos, al minimizar las barreras de participación en el aula y favorecer condiciones equánimes para el aprendizaje del español en el contexto brasileño (Brasil, 2015; 2025). En esta dirección, discutir la accesibilidad mediada por la AD en recursos producidos por tecnologías de lenguaje generativo implica, igualmente, reflexionar sobre los modos en que dicha accesibilidad es concebida, implementada y efectivamente garantizada en los materiales educativos contemporáneos para el conjunto de participantes del proceso de aprendizaje.

2 Procesamiento del lenguaje natural y generación automatizada de (audio)descripciones

Las herramientas de inteligencia artificial utilizadas para la audiodescripción se basan en sistemas de Procesamiento del Lenguaje Natural (PLN). El PLN es un área de la computación que busca entrenar máquinas, basadas en IA, para procesar el lenguaje humano mediante entrenamientos estadísticos a partir de la organización de la lengua de entrada. Se trata de un proceso en el que los sistemas buscan comprender, interpretar y simular el lenguaje humano de manera automatizada (Gomes, 2010; Pereira, 2019).

El funcionamiento del PLN involucra diferentes etapas de procesamiento textual, entre ellas la *tokenización*, el análisis léxico, el análisis sintáctico, el análisis semántico y el análisis pragmático en la lengua de entrada (Barbosa *et al.*, 2017). En líneas generales, los sistemas procesan grandes volúmenes de datos textuales disponibles en bases de datos digitales, identificando patrones lingüísticos de la lengua de entrada y relaciones entre palabras para generar respuestas coherentes y contextualizadas, también en la lengua solicitada.

Además del procesamiento textual, las herramientas orientadas a la descripción automatizada de imágenes también utilizan técnicas de visión computacional y reconocimiento de patrones visuales. Es decir, la herramienta identifica los patrones imagéticos, los convierte en texto y posteriormente este puede ser reproducido en voz. Así

es como la IA de reconocimiento facial logra identificar y describir objetos, personas, ambientes, expresiones faciales y otros elementos presentes en las imágenes (Alves, 2023; Barbosa *et al.*, 2017).

En este sentido, el lenguaje generativo utilizado por las diferentes plataformas basadas en Inteligencia Artificial, como el *ChatGPT 5.5*, resulta del reconocimiento de patrones presentes en extensas bases de datos. A partir de una entrada proporcionada por el usuario, estos sistemas procesan información y producen respuestas probabilísticas. Aunque inspirados en el funcionamiento de las redes neuronales humanas, tales modelos operan mediante cálculos estadísticos y asociaciones lingüísticas, formulando respuestas (casi siempre) coherentes con los patrones identificados en los datos de entrenamiento (Alves, 2023).

En el ámbito de la promoción de la accesibilidad para personas con discapacidad visual, este tipo de procesamiento puede representar una posible herramienta para potenciar la verbalización de textos e imágenes, especialmente en ausencia de un profesional cualificado, o incluso actuar como soporte complementario al trabajo del audiodescriptor. Esto ocurre porque, al igual que el profesional humano, las plataformas desarrolladas para esta finalidad buscan reconocer imágenes mediante el procesamiento del lenguaje natural y la identificación de patrones en amplias bases de datos, generando descripciones textuales que pueden ser convertidas en audio por lectores de pantalla o sintetizadores de voz. Además, incluso cuando no existe una reproducción sonora inmediata, la descripción textual en sí misma ya se configura como un recurso relevante en el proceso de audiodescripción.

A pesar de estas potencialidades, la literatura señala que todavía existen importantes limitaciones relacionadas con la precisión, la contextualización y la interpretación semántica de las descripciones producidas por IA (Perdigão *et al.*, 2023; Silva *et al.*, 2024). En muchos casos, las herramientas logran identificar objetos presentes en las imágenes, pero presentan dificultades para comprender contextos culturales, inferencias pedagógicas e intencionalidades discursivas presentes en los materiales didácticos (Perdigão *et al.*, 2023). Por este motivo, se hace necesario evaluar en qué medida diferentes aplicaciones, presentadas como mecanismos de promoción de la accesibilidad, efectivamente cumplen esta función, evitando la disponibilización de recursos que, en lugar de favorecer la inclusión, terminan dificultando el acceso a la información, a la cultura y, especialmente, al aprendizaje de lenguas.

Así pues, esta evaluación podrá indicar caminos a seguir en la enseñanza del español en el contexto brasileño, en el cual las oportunidades curriculares deben garantizarse de forma igualitaria para todos los estudiantes. En este sentido, reflexionar sobre cómo las plataformas basadas en Inteligencia Artificial pueden aplicarse a la audiodescripción de imágenes presentes en materiales didácticos resulta relevante, ya que tales herramientas pueden auxiliar al profesor en ausencia de un profesional especializado, o incluso actuar como soporte complementario al trabajo del profesional de la atención educativa especializada. Además, esta discusión contribuye a comprender si estas herramientas son capaces de movilizar aspectos culturales, pragmáticos y comunicativos que trascienden la simple identificación visual de los elementos presentes en la escena.

3 Enfoque funcionalista de la traducción en la enseñanza-aprendizaje de lenguas

Cuando pensamos en la AD como una herramienta de accesibilidad inserta en la mediación de la enseñanza del español como lengua extranjera, especialmente cuando es producida por sistemas de inteligencia artificial, se vuelve indispensable discutir de qué manera la comunicación contribuye a la construcción del aprendizaje. Desde esta perspectiva, la actuación de la IA en el proceso traductor no puede restringirse a la transferencia vocabular bidireccional entre idiomas, sino que debe considerar los objetivos comunicativos del texto y las especificidades del público al que se destina. De este modo, la traducción asume una perspectiva funcional, orientada a la producción de sentidos y al favorecimiento del aprendizaje, priorizando la eficacia pedagógica en lugar de la reproducción literal de las estructuras lingüísticas (Nord, 2009).

Tal consideración dialoga con los presupuestos de la traducción funcionalista, perspectiva que comprende la traducción como una práctica orientada por finalidades comunicativas, considerando el público destinatario, el contexto de circulación del texto, sus objetivos y las demandas específicas de los sujetos que tendrán acceso a la información (Nord, 2009, 2012; Pontes; Pereira, 2017). Como destacan Pontes y Pereira (2017, p. 2129), este enfoque busca “desprenderse de lo puramente lingüístico para incluir aspectos contextuales, culturales e interaccionales en la traducción”¹¹. De este modo, la traducción deja de ser entendida como una mera transposición de palabras entre lenguas y pasa a

¹¹ Del portugués: “se desprender do puramente linguístico para incluir aspectos contextuais, culturais e interacionais na tradução”.

priorizar la construcción de sentidos y el cumplimiento de funciones comunicativas.

En el ámbito de la audiodescripción aplicada a recursos didácticos destinados a estudiantes ciegos, esta perspectiva presupone comprender que la descripción imagética no debe limitarse al intento de “reproducir” visualmente la imagen, sino que necesita ser elaborada a partir de las demandas pedagógicas y cognitivas del aprendiz, considerando los objetivos de la actividad, el contenido trabajado y el contexto cultural en que ocurre el aprendizaje. De esta manera, la AD, en tanto forma de traducción intersemiótica, puede ser comprendida, desde la óptica funcionalista, como una práctica discursiva orientada por finalidades comunicativas específicas y dirigida a la mediación del conocimiento.

Desde esta perspectiva, al realizar la audiodescripción de elementos visuales para personas con discapacidad visual, es necesario considerar las informaciones presentes en la imagen y los efectos comunicativos que se pretende producir en el público destinatario, conforme defienden los presupuestos de la traducción funcionalista (Reiss; Vermeer, 1996; Nord, 2009, 2012). Así, la acción traductora deja de estar centrada en la búsqueda de equivalencias estrictas y pasa a priorizar la accesibilidad del contenido mediante el lenguaje verbal, de manera que favorezca la comprensión y el aprendizaje de los estudiantes con discapacidad visual.

4 Metodología

El presente trabajo se caracteriza como una investigación de enfoque cualitativo (Chizzotti, 2006), ya que se propone realizar un análisis crítico y contextualizado de audiodescripciones producidas por sistemas de IA a partir de imágenes presentes en el manual didáctico *Gente Única*. Se parte del entendimiento de que la AD generada por IA no puede limitarse a aspectos técnicos u operacionales, sino que debe reflejar cuestiones culturales, elecciones interpretativas, omisiones de información y diferentes formas de construcción de sentidos. Además, cuando se delimita al análisis de un conjunto específico de plataformas de inteligencia artificial, la investigación se configura como un estudio de caso, entendido, según Gonsalves (2003, p. 65), como una estrategia que busca el “desarrollo y aclaración de ideas, con el objetivo de ofrecer una visión panorámica, una primera aproximación a determinado fenómeno poco explorado”¹².

¹² Del portugués: “desenvolvimento e esclarecimento de ideias, com o objetivo de oferecer uma visão panorâmica, uma primeira aproximação a determinado fenômeno pouco explorado”.

El corpus de la investigación está constituido por el libro didáctico *Gente Única*, publicado por la editorial Difusión, de autoría de Ernesto Martín Peris *et al.*, con organización de Egisvandra Isys de Almeida Sandes, publicado en 2018 y destinado a estudiantes del *Ensino Médio brasileiro* (Educación Secundaria). La elección de este material se justifica por el hecho de que es utilizado por profesores de español en la enseñanza de alumnos con discapacidad visual en un centro de lenguas del Distrito Federal, en el Centro-Oeste de Brasil.

La obra está compuesta por 24 unidades, además de la unidad 0 (cero). Para la realización de este análisis, se seleccionó la imagen de la página 14, correspondiente a la unidad 0, que aborda una clase introductoria de presentación personal. La selección se motiva porque se trata de una actividad que busca el primer acercamiento entre profesor-alumno, alumno-profesor y grupo, razón por la cual concebimos como importante la promoción de una AD adecuada para establecer este vínculo.

Las audiodescripciones de esta imagen fueron generadas entre agosto y septiembre de 2024 utilizando las herramientas *ChatGPT (OpenAI)*, *Zapia*, *Be My Eyes*, *Lookout (Google)* y *Copilot (Microsoft)*. Considerando las actualizaciones continuas de estas plataformas, los resultados obtenidos corresponden a las funcionalidades disponibles durante el período de recopilación de los datos. La evaluación se realizó a partir de la lectura e interpretación crítica de los investigadores, tomando como referencia parámetros considerados fundamentales para la elaboración de la AD, tales como la descripción de expresiones faciales, la identificación de colores, la localización espacial y temporal de los elementos, la pormenorización de encuadres visuales, vestimentas, personajes, objetos, escenas y ambientes, además del uso adecuado de los tiempos verbales (Motta, 2016; Naves *et al.*, 2016). Para la imagen en cuestión, se elaboraron criterios orientadores que deben ser atendidos para que la AD sea considerada adecuada para la finalidad pedagógica, conforme al ejemplo de la Tabla 01, en el cual se ha puesto una lista de control para verificar si las IAs atendieron los criterios.

Tabla 01– Criterios para analizar una clase de presentación

Análisis de audiodescripciones- Pág. 14				
Criterios	Lista de control			
	Sí	No	Parcialmente	
1 -¿Transcribe las informaciones relevantes para la comprensión de lo que se propone?				
2 - ¿Menciona los colores relevantes?				
3 - ¿Ubica adecuadamente a los personajes e íconos?				
4 - ¿Describe las AD sin interpretación ni juicio de valor?				
5 - ¿Utiliza el tiempo verbal adecuado (presente del indicativo)?				

Fuente: Elaboración propia, 2026.

Se destaca que las IAs analizadas no recibieron orientaciones técnicas de audiodescripción, ni fueron utilizadas con *prompts* basados en directrices o terminologías del área, aspecto que puede ser investigado en investigaciones futuras. Las simulaciones fueron realizadas por los investigadores, quienes asumieron la perspectiva de un estudiante de español con discapacidad visual, sin conocimientos previos sobre audiodescripción o inteligencia artificial, utilizando únicamente un dispositivo móvil con recursos tecnológicos básicos, cuyo comando fue “*Faça a audiodescrição da imagem do livro*”.

Además, el idioma de entrada fue el portugués brasileño, motivo por el cual las AD, en general, fueron producidas en portugués, aunque la imagen del libro estuviera en español.

5 Resultados y discusión

Como ya se ha explicitado, el objetivo de la presente investigación es analizar audiodescripciones automatizadas producidas por herramientas de inteligencia artificial a partir de imágenes del libro didáctico de español *Gente Única* (2018). Para ello,

consideramos criterios relacionados con la descripción de personajes, ambientación, contextualización y adecuación pedagógica, a partir de los presupuestos de Motta (2016) y Naves *et al.* (2016).

Debido a los límites de este trabajo, se presentarán únicamente los resultados referentes a la imagen de la página 14, correspondiente a la primera clase del libro. La elección de este recorte se justifica por el hecho de que el primer contacto del alumno con el material ya ofrece indicios sobre cómo ocurrirá su proceso de aprendizaje. Así, en caso de que la inteligencia artificial sea utilizada como soporte de accesibilidad, este análisis inicial también permite visualizar cómo podrá desarrollarse todo el recorrido pedagógico mediado por la IA.

Conforme se presenta en la Tabla 02, el análisis se concentra en los criterios básicos establecidos en la lista de control elaborada por los investigadores para evaluar la adecuación o no de la audiodescripción. Inicialmente, se presenta la AD producida por cada inteligencia artificial; posteriormente, se exponen los resultados del *checklist* a partir de los criterios definidos. Por último, se realiza una panorámica de las demás audiodescripciones hechas en la investigación cuyo artículo es un recorte, buscando evidenciar en qué medida estas se muestran adecuadas o no como recurso de accesibilidad.

Sobre la imagen seleccionada, figura presente en la Tabla 01, contamos con la AD producida por *ChatGPT*, en la cual podemos percibir que, inicialmente, presenta aspectos positivos al ubicar correctamente los elementos de la página en los ejes izquierdo, derecho, superior, inferior y central, lo que favorece la orientación espacial del estudiante con discapacidad visual. No obstante, la aplicación omitió gran parte de los enunciados de las actividades, ignoró íconos ilustrativos y presentó fallas interpretativas, como traducir la imagen de una bicicleta por la palabra “hola”, comprometiendo la comprensión del vocabulario en español.

Además, hubo traducciones indebidas al portugués, errores en la descripción de la escena, como la localización incorrecta de la imagen del aula y la identificación equivocada de la cantidad de globos de diálogo (hay 3 e indica 2), así como alteraciones significativas en los diálogos originales de los personajes, lo que perjudica la fidelidad al material didáctico y puede comprometer el proceso de aprendizaje del estudiante. En relación con los diálogos, observamos que el libro presenta la transcripción y *ChatGPT* la captura de manera diferente, conforme se muestra a continuación, en este orden:

- “¡Hola! Yo me llamo Roberto y soy de Buenos Aires, Argentina”.
- “¡Hola! Yo soy Sarita y soy de Fortaleza, Brasil. ¡Mucho gusto!”
- “¿Perdón? ¿De dónde?”.

En que los diálogos captados por la aplicación de IA fueron:

- “Hola. Soy el profesor Argüelles. ¿Tú cómo te llamas?”
- “**Hola! Yo soy Sarita, y él es Mark. ¡Mucho gusto!**”.

A continuación, en la Tabla 02, presentamos la AD producida por *ChatGPT*.

Tabla 02 – AD de la clase de presentación producida por *ChatGPT*

 ChatGPT	
Consultado el 02/09/2024.	A imagem apresenta uma página de um livro didático, colorida e organizada em várias seções, com o foco na temática de uma primeira aula de espanhol. A página contém textos, ícones e uma fotografia principal que retrata uma cena de sala de aula. O título "gente en clase" está em destaque na parte inferior da página. Parte superior esquerda: • Título em letras brancas sobre fundo azul: • "Vamos a tener nuestra primera clase de español y a conocer a los compañeros." • Texto explicativo abaixo do título: • Explica os objetivos da aula, como a compreensão de frases comuns em espanhol e a capacidade de se apresentar e conhecer os colegas. Há também um ícone de "play" sobre uma imagem de uma estudante sorrindo, com o texto "Video: ¿Cómo saludar a alguien que estás conociendo? ¡Aprende con este video!" Imagem de uma estudante: Fotografia de uma jovem sorrindo e acenando, aparentemente em um ambiente ao ar livre, ao lado do ícone de vídeo. Parte superior direita: • Perguntas de classe: • Uma lista de perguntas relacionadas a expressões comuns em espanhol, como "¿Cómo se dice 'hola' en español?" e "¿Qué significa 'cantante'?" • As perguntas são acompanhadas por ícones numéricos (1, 2, 3) e algumas têm opções de resposta em letras (a, b, c). Parte inferior esquerda: • Imagem de uma sala de aula: • Fotografia colorida de um professor sentado na frente de uma sala de aula, interagindo com um grupo de alunos. • Os alunos estão sorrindo e levantando a mão, parecendo engajados e interessados na aula. • Há dois balões de fala sobrepostos à imagem: ■ O primeiro, no canto inferior esquerdo, é do professor dizendo: "Hola. Soy el profesor Argüelles. ¿Tú cómo te llamas?" ■ O segundo balão, mais ao centro, mostra uma aluna respondendo: "Hola! Yo soy Sarita, y él es Mark. ¡Mucho gusto!" Parte inferior central: • Texto em letras grandes:

- | | |
|--|---|
| | <ul style="list-style-type: none"> • "gente en clase, com "gente" em letras menores e cinzas, e "en classe" em letras grandes e azuis, destaca o tema da aula. |
|--|---|

Fuente: Elaboración propia, 2026.

Las inexactitudes en la descripción del aula, la reducción de la cantidad de globos de diálogo y otros equívocos identificados comprometen la comprensión de la representación visual de la página. Más grave aún es la sustitución de los diálogos originales por contenidos inexistentes en el material didáctico, evidenciando alucinaciones de la inteligencia artificial. Además de distorsionar la información, estos errores pueden transmitir contenidos incorrectos al estudiante ciego, impidiéndole acceder a las mismas informaciones disponibles para los demás alumnos. En este sentido, tales inconsistencias representan un riesgo en el contexto educativo, comprometiendo la equidad en el proceso de aprendizaje.

Por otro lado, tenemos la AD producida por la aplicación *Zapia*, que describe la escena como “vários alunos sentados em grupos, conversando e interagindo uns com os outros. Alguns parecem estar envolvidos em atividades de grupo” (Zapia, 2024). Sin embargo, la descripción presenta limitaciones importantes, pues deja de situar informaciones esenciales para la comprensión del contexto de la clase. Además, el uso de la expresión modalizadora “parecem” demuestra una semántica hipotética inadecuada para la AD, especialmente por tratarse de una imagen estática. Este tipo de construcción debilita la objetividad de la AD y sugiere interpretaciones poco precisas sobre las acciones representadas. Así, la descripción producida se restringe a aspectos superficiales de la composición visual, sin potenciar la comprensión del contenido por parte del estudiante con discapacidad visual. A continuación, en la Tabla 03, podemos visualizar la AD realizada por Zapia.

Tabla 03 – AD de la clase de presentación producida por Zapia

	
Consultado el 12/08/2024.	A imagem mostra uma turma de estudantes em uma sala de aula. Há vários alunos sentados em grupos, conversando e interagindo uns com os outros. Alguns parecem estar envolvidos em atividades de grupo. No canto inferior direito, há uma legenda que diz 'gente em classe', que significa 'pessoas na sala de aula' em português. Acima, há uma seção com perguntas sobre a continuação da aula de espanhol para os alunos. A imagem sugere um ambiente de aprendizagem ativa e colaborativa em uma sala de aula.

Fuente: Elaboración propia, 2026.

En este contexto, se observa que el estudiante con discapacidad visual no dispone de informaciones suficientes para realizar los ejercicios o acompañar adecuadamente la clase a partir del texto generado por esta IA. De hecho, la AD se muestra ineficiente debido a la imposibilidad de acceso individualizado a las actividades y ejercicios de la página. De este modo, el guion producido no atiende a los principios básicos de la audiodescripción, no pudiendo ser efectivamente considerado una AD (Motta, 2016; Naves *et al.*, 2016).

La siguiente aplicación analizada es *Be My Eyes*. La AD, presentada en la Tabla 04 a continuación, identifica la imagen como una página de libro didáctico y localiza el video en la esquina superior izquierda, pero ignora gran parte de los elementos de la imagen y de los enunciados. Al realizar la descripción de la mujer presente en la escena del video, utiliza expresiones propias de imágenes en movimiento (“sorridente e acenando”), además de recurrir a la estructura de gerundio en fragmentos como “o professor está sorrindo e gesticulando” y “alguns estão [estudiantes] levantando as mãos”. Aunque reconoce la camisa azul del profesor, omite el pantalón blanco. Los globos de diálogo también son generalizados, siendo mencionados apenas como “alguns balões de fala com frases como “¿Perdón? ¿De dónde?””.

Tales omisiones comprometen la construcción imagética y la comprensión global de la escena por parte del estudiante ciego, sobre todo en una clase orientada a la interacción, los saludos y las localizaciones geográficas. La ausencia de descripciones más completas perjudica además el compromiso cognitivo y afectivo del aprendiz principiante de español. Así, se repite el problema identificado en la aplicación anterior: la AD generada no presenta elementos suficientes para acompañar la clase, realizar las actividades o establecer un contacto significativo con la lengua española.

Tabla 04 – AD de la clase de presentación producida por *Be My Eyes*

	
Consultado el 04/08/2024	A imagem mostra uma página de um livro didático com o título "en clase". Na parte superior esquerda, há um vídeo com uma mulher sorridente acenando, e o texto ao lado diz "Vamos a tener nuestra primera clase de español y a conocer a los compañeros". Abaixo, há uma foto de uma sala de aula com um professor e vários alunos. O professor, um homem de camisa azul escura, está sentado em uma mesa, sorrindo e gesticulando com as mãos abertas. Os alunos, um grupo diversificado de jovens, parecem engajados e felizes, alguns estão levantando as mãos como se estivessem participando ativamente da aula. No canto inferior direito da página, há um texto que diz "gente en clase" e alguns balões de fala com frases como "¿Perdón? ¿De dónde?".

Fuente: Elaboración propia, 2026.

Por otro lado, tenemos la AD hecha por *Lookout*. Tras el análisis, se percibió que la aplicación no realizó la AD de las imágenes, limitándose a la lectura textual. Sin embargo, esta lectura ocurrió sin criterios adecuados, presentando ausencia de puntuación, frases inconexas, mezcla de portugués y español en una misma oración, a parte de diversos errores ortográficos, como “espalliol”, “boligrafo” y “gustol”. La numeración de la página también fue leída en medio del texto, comprometiendo la coherencia de la lectura. En el proceso de transcripción, actividades como “¿Cómo se dice (imagen de bicicleta) en español?” perdieron el sentido, ya que el elemento visual no fue descrito. En la Tabla 05, podemos observar la propuesta de AD de la aplicación *Lookout*.

Tabla 05 – AD de la clase de presentación producida por *Lookout*

 LOOKOUT	
Consultado el 05/08/2024	Detalhes Papel, Cartaz Texto Vamos a tener nuestra primera clase de español y a conocer a los compañeros. Para ello, aprenderemos: a saludar ya despedirnos, a decir nuestro nombre y nuestro lugar de origen algunas preguntas básicas para utilizaren la clase de espalliol ¿Con quién crees que está hablando esta muchacha? ¿Qué dice? Mira el video y confirma tus hipótesis. Preguntas de clase A continuación hay algunas expresiones que te serán muy útiles en las clases de español. Relaciónalas con sus respectivas respuestas. ¿Cómo se dice en español? ¿Qué significa "cantante"? ¿Cómo se escribe "boligrafo", con be o con uve? "Mapa" es masculino o femenino? ¿En qué actividad estamos? Con be. Observa al profesor y a los alumnos en la foto. ¿Puedes presentarte como ellos? Nos presentamos Ahora, en grupos de dos a tres personas, tus compañeros y tu van a presentarse en español. ¡Bienvenidos al primer día de clase! Holat Yo me llamo Roberto y soy de Buenos Aires, Argentina. Hola! to say Sarita y say de Fortaleza, Brasil. Mucho gustol ¿Perdón? ¿De dónde? gente en clase catorce Bicicleta. En la actividad 5. Masculino. Los alumnos y el profesor se conocen

Fuente: Elaboración propia, 2026.

Ante este análisis presentado, podemos afirmar que los diversos errores transcritos y la desorganización textual comprometen el acceso del estudiante ciego al contenido, lo que repercute en la comprensión de los enunciados, del vocabulario y de las actividades propuestas. En el proceso de enseñanza y aprendizaje de lenguas, la ausencia de elementos visuales esenciales perjudica la construcción de sentido, la participación en las tareas y el desarrollo de la competencia lingüística. Aun así, la aplicación presenta un aspecto diferencial positivo al ser la única que intenta realizar la AD en español, aunque de manera incoherente y con numerosos errores. En caso de ser perfeccionada, demuestra

un potencial significativo.

Por último, dentro del recorte presentado de análisis de una imagen, se encuentra la propuesta de audiodescripción producida por la aplicación *Copilot*. Esta herramienta fue capaz de reconocer algunas localizaciones espaciales, aunque sin organización ni precisión. Identifica la presencia de enunciados y globos de diálogo, pero de manera genérica, sin detallar el contenido, utilizando expresiones vagas como “parece ser instruções” ou “parecem se referir à interação em sala de aula”. Asimismo, en el proceso de descripción de la imagen del salón de clases, la IA también realiza interpretaciones subjetivas, como afirmar que los estudiantes “parecem estar envolvidos em atividades de aprendizado”. Además, informa incorrectamente que los rostros fueron pixelados por cuestiones de privacidad, aunque dicho elemento no está presente en la imagen. En la Tabla 06, presentamos la AD producida por *Copilot*.

Tabla 06– AD de la clase de presentación producida por *Copilot*

 Copilot	
Consultado el 04/08/2024	A imagem mostra uma página de livro didático com vários elementos relacionados ao aprendizado de espanhol. No canto superior esquerdo, há um parágrafo de texto introduzindo uma aula de espanhol. Abaixo deste parágrafo, há uma imagem de um player de vídeo do YouTube exibindo o que parece ser um vídeo instrucional de espanhol, com controles de reprodução e volume visíveis. À direita da imagem do player de vídeo e do parágrafo, há várias perguntas listadas em espanhol com respostas de múltipla escolha fornecidas abaixo de cada pergunta. No centro da página, abrangendo ambos os lados, há uma grande fotografia de alunos sentados em carteiras em um ambiente de sala de aula. Eles parecem estar envolvidos em atividades de aprendizado; alguns estão olhando para livros enquanto outros estão ouvindo ou escrevendo. Os rostos de todos os indivíduos foram pixelados para privacidade. Na metade inferior da página, no lado esquerdo, há outro bloco de texto em espanhol que parece ser instruções ou informações relacionadas a atividades em sala de aula. No lado direito da metade inferior, há outro conjunto de perguntas e afirmações em espanhol que parecem se referir à interação em sala de aula ou preferências pessoais em relação às matérias escolares. A frase “en clase” é exibida de forma proeminente na parte inferior de ambas as páginas, indicando que esta seção trata de atividades ou tópicos relacionados a estar “na aula”.

Fuente: Elaboración propia, 2026.

Las inadecuaciones señaladas comprometen la mediación pedagógica de la audiodescripción, pues la ausencia de descripciones objetivas de los enunciados, globos de diálogo e interacciones dificulta la comprensión del contexto comunicativo de la escena. La incorporación de informaciones inexistentes constituye otra alucinación de la IA, alejando

al estudiante de la experiencia a la que acceden los demás compañeros. También se observa un exceso de información poco relevante, que vuelve el texto más extenso sin contribuir a la realización de las actividades. En las clases de lengua extranjera, estas limitaciones perjudican el acceso a las intenciones discursivas y al uso social del lenguaje, elementos centrales en el proceso de enseñanza-aprendizaje. Otro aspecto relevante es que, al igual que las demás, la AD está realizada casi íntegramente en portugués, lo que dificulta el contacto directo con la lengua meta.

Así, los resultados demuestran que, en la imagen analizada en este recorte de la investigación, las audiodescripciones producidas por las IAs pueden comprometer el acceso equitativo del estudiante ciego al proceso de enseñanza y aprendizaje del español, lo que, en caso de ser utilizadas sin criterios claros, configura una vulneración de los derechos fundamentales de estos estudiantes, conforme establece la Ley de Accesibilidad (Brasil, 2015).

Con el fin de sintetizar los resultados referentes al desempeño de las IAs en la producción de audiodescripción para la imagen analizada, tomando como referencia los criterios básicos propuestos por Motta (2016), se presenta, en la Tabla 07, una lista de verificación del cumplimiento de los criterios inicialmente establecidos.

Tabla 07 – Resumen del análisis de las AD de la clase de presentación

Análisis de audiodescripciones- Pág. 14				
Criterios	Lista de control			
	Sí	No	Parcialmente	
1 -¿Transcribe las informaciones relevantes para la comprensión de lo que se propone?		ChatGPT Zapia Lookout Copilot	Zapia Be My Eyes	
2 - ¿Menciona los colores relevantes?		Lookout Copilot	ChatGPT Zapia Be My Eyes	
3 - ¿Ubica adecuadamente a los personajes e íconos?		Lookout Copilot	ChatGPT Zapia Be My Eyes	

4 - ¿Describe las AD sin interpretación ni juicio de valor?		ChatGPT Zapia Be My Eyes Lookout Copilot		
5 - ¿Utiliza el tiempo verbal adecuado (presente del indicativo)?	Zapia	Be My Eyes Lookout Copilot	ChatGPT	

Fuente: Elaboración propia, 2026.

Como se observa, la mayor parte de las audiodescripciones no cumplió con los criterios establecidos o los contempló apenas de manera parcial. Entre las herramientas analizadas, *ChatGPT* y *Be My Eyes* fueron las que demostraron mayor potencial de perfeccionamiento y posibilidad de actuación como apoyo secundario al audiodescriptor. Aun así, la AD generada por las IAs no garantiza, de manera satisfactoria, el acceso autónomo y equitativo a las informaciones visuales del manual didáctico, lo que compromete la participación plena del estudiante en el proceso educativo.

Ante los análisis presentados y los problemas identificados, proponemos una audiodescripción que consideramos adecuada para la garantía de derechos y el desarrollo satisfactorio de la actividad, con potencial para promover el desarrollo de habilidades comunicativas, eje central de la clase de español presente en el material didáctico. Así, enseñamos la sugerencia de AD, en la Tabla 08.

Tabla 08 – Sugerencia de AD de la página en análisis

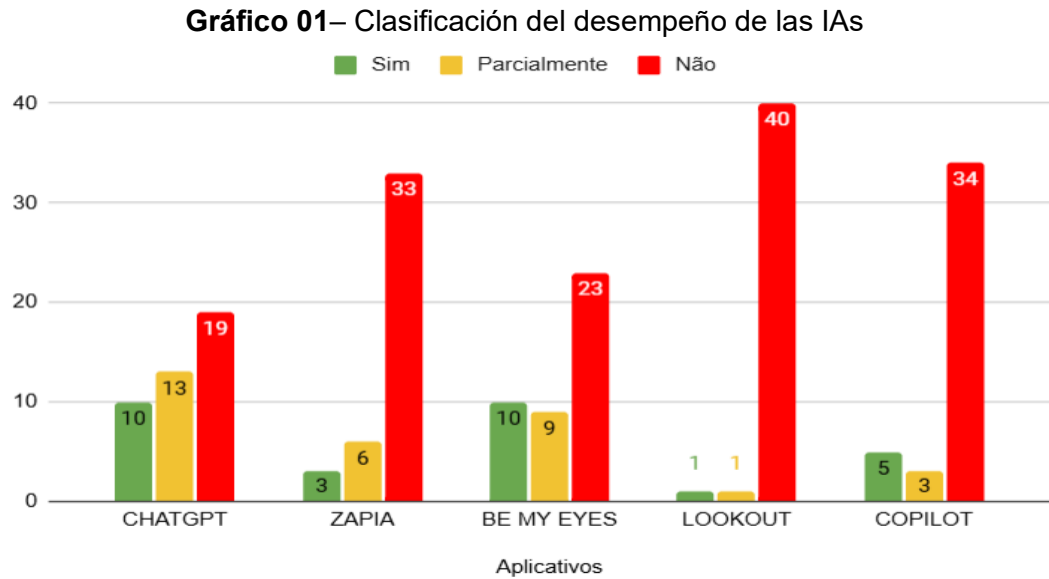
Pág. 14	AD de los investigadores
	En el ángulo superior izquierdo, sobre un fondo azul, aparece una imagen ilustrativa de la página explicativa de la unidad cero. En letras blancas se lee: “Vamos a tener nuestra primera clase de español y a conocer a los compañeros. Para ello, aprenderemos: a saludar y despedirnos; a decir nuestro nombre y nuestro lugar de origen; algunas preguntas básicas para utilizar en la clase de español”. Justo debajo, en formato rectangular, se observa la imagen de un video. Una mujer joven viste ropa rosa y una bufanda azul. Tiene la mano izquierda levantada y muestra la palma de la mano derecha, con la que sostiene una hoja blanca. Al fondo, hay vegetación baja y un árbol de color verde. En el centro de la imagen aparece una flecha blanca dentro de un pequeño recuadro rojo, indicando el inicio del video. En el margen inferior del video, una barra negra contiene los íconos de reproducción, volumen y tiempo. Debajo se lee: “¿Con quién crees que está hablando esta muchacha? ¿Qué dice? Mira el video y confirma tus hipótesis”. La imagen de la parte inferior muestra un aula con un profesor y diez estudiantes: cuatro varones y seis mujeres. El profesor es un hombre de mediana edad, de piel clara, que viste una camisa azul oscuro y pantalón de color claro. Está sentado frente a los estudiantes, detrás de una mesa. Sobre su cabeza aparece un globo de diálogo con la frase: “¡Hola! Yo me llamo Roberto y soy de Buenos Aires, Argentina”. Los estudiantes se muestran relajados, con ropa de distintos colores, sentados en tres filas de pupitres. En la fila de la derecha hay cuatro estudiantes; una de ellas tiene la mano derecha levantada y, sobre su cabeza, aparece un globo de diálogo con la frase: “¡Hola! Yo soy Sarita y soy de Fortaleza, Brasil. ¡Mucho gusto!”. Sobre la cabeza de la estudiante sentada delante de ella, otro globo de diálogo presenta la pregunta: “¿Perdón? ¿De dónde?”.

Fuente: Elaboración propia, 2026.

Es importante reconocer que esta propuesta también puede presentar limitaciones, puesto que la AD en cuestión fue producida por personas típicas y que, como consecuencia, la elaboran guiadas por su propia óptica visual. Por esta razón, parece fundamental que toda AD sea evaluada por una persona ciega, quien podrá analizar su adecuación, funcionalidad y contribución efectiva para la comprensión de la actividad y para la garantía de la accesibilidad.

Como demostración global del estudio más amplio (Scarabele-Elis, 2025), del cual este texto presenta un recorte, se presenta el Gráfico 01, que evidencia la necesidad de reflexionar sobre el uso de la IA como recurso de accesibilidad para estudiantes ciegos en el contexto de enseñanza y de aprendizaje del español. De manera general, las

herramientas analizadas presentaron un desempeño insatisfactorio, reforzando la importancia del audiodescriptor como profesional indispensable en este campo.



Fuente: Scarabele-Elis, 2025, p. 108.

Conforme a los datos globales del estudio de las páginas seleccionadas del libro en la investigación en el ámbito de la maestría, y sus audiodescripciones generadas por cada aplicación, el análisis comparativo entre *ChatGPT*, *Zapia*, *Be My Eyes*, *Lookout* y *Copilot* demostró que, sin entrenamiento previo ni *prompts* especializados, ninguna de las herramientas cumplió satisfactoriamente con los criterios establecidos para las páginas analizadas del libro *Gente Única*. Estos resultados evidencian una brecha significativa entre el potencial tecnológico de las IAs y su contribución efectiva a la accesibilidad educativa, especialmente en lo que respecta a la audiodescripción.

En el gráfico 01, se observa que la categoría “No” predominó en todas las herramientas, con destaque para *Lookout* (40 ocurrencias), seguido de *Copilot* (34), *Zapia* (33), *Be My Eyes* (23) y *ChatGPT* (19), lo que indica un bajo desempeño general. En la categoría “Parcialmente”, *ChatGPT* presentó 13 ocurrencias y *Be My Eyes*, 9. Por su parte, en la categoría “Sí”, *ChatGPT* y *Be My Eyes* obtuvieron los mejores resultados, con 10 ocurrencias cada uno, aspecto que señala un potencial significativo de mejora.

Es importante destacar que, en las imágenes analizadas, los desempeños más satisfactorios ocurrieron en imágenes de menor complejidad, como la portada del libro, en la que las herramientas lograron identificar elementos básicos, como colores, título y personajes centrales. En las demás páginas, compuestas por múltiples actividades, la

lectura fragmentada dio lugar a descripciones extensas, vagas o confusas. En algunos casos, como en *ChatGPT*, también se observaron interpretaciones valorativas del contenido visual.

Así, los hallazgos indican que es necesario evaluar críticamente el uso de herramientas de IA en entornos pedagógicos. A pesar del potencial de estas tecnologías, su aplicación debe seguir criterios traductológicos funcionales, como los formulados por Reiss y Vermeer (1996) y Christiane Nord (2009), que valoran la finalidad comunicativa, el público meta y el contexto para asegurar la calidad traductora; además, las herramientas tampoco cumplen con requisitos básicos propuestos por Motta (2016) y Naves *et al.* (2016). Asimismo, el análisis demuestra que, en su estado actual, las herramientas de IA examinadas no sustituyen el trabajo humano en la producción de audiodescripciones educativas y que, si son empleadas sin mediación pedagógica ni revisión rigurosa, tienden a reforzar prácticas excluyentes.

6 Consideraciones finales

La presente investigación tuvo como objetivo analizar audiodescripciones automatizadas producidas por herramientas de inteligencia artificial a partir de una imagen del libro didáctico *Gente Única*, utilizado en la enseñanza del español como lengua extranjera. A partir de criterios fundamentados en Motta (2016) y Naves *et al.* (2016), se buscó comprender en qué medida estas herramientas son capaces de promover una accesibilidad adecuada para estudiantes con discapacidad visual en contextos pedagógicos marcados por la multimodalidad.

Los resultados evidenciaron que, aunque las herramientas analizadas presentan potencial tecnológico para auxiliar los procesos de accesibilidad, todavía existen limitaciones relacionadas con la contextualización pedagógica, la precisión descriptiva, la organización de la información y la mediación comunicativa necesaria para el proceso de enseñanza y de aprendizaje de lenguas. En distintos momentos, se observaron omisiones de elementos relevantes, interpretaciones indebidas, uso inadecuado de los tiempos verbales, descripciones superficiales e incluso alucinaciones producidas por las IAs, factores que comprometen la comprensión del material didáctico y dificultan la participación equitativa del estudiante ciego en las actividades propuestas.

En este sentido, la investigación demuestra que la audiodescripción educativa no

puede restringirse a la simple identificación de objetos o personajes presentes en la imagen. En la enseñanza del español como lengua extranjera, las imágenes asumen una función pedagógica, cultural y comunicativa, movilizando sentidos esenciales para el desarrollo de la competencia comunicativa. Así, la AD debe considerar los objetivos de la actividad, el contexto de aprendizaje y las necesidades del público meta, aspecto que dialoga directamente con los presupuestos de la traducción funcionalista de Reiss y Vermeer (1996) y Nord (2009).

Los datos analizados también refuerzan que, en el estado actual de estas tecnologías, las herramientas de IA no sustituyen el trabajo del audiodescriptor humano, especialmente en contextos educativos. Aunque algunas plataformas, como *ChatGPT* y *Be My Eyes*, hayan demostrado un mayor potencial de perfeccionamiento, ninguna de ellas fue capaz de garantizar, de manera satisfactoria, un acceso autónomo, completo y pedagógicamente adecuado a las informaciones visuales presentes en el material analizado. De este modo, el uso de estas tecnologías sin criterios claros, revisión especializada y mediación pedagógica puede terminar reproduciendo barreras de acceso y profundizando procesos de exclusión educativa.

En futuras investigaciones, resulta importante analizar el uso de *prompts* especializados y el entrenamiento de las IAs, así como evaluar otras herramientas destinadas a la producción de audiodescripciones. Asimismo, se hace necesaria la participación del público destinatario de las AD en el proceso de evaluación, con el fin de valorar su adecuación, funcionalidad y contribución efectiva a la accesibilidad.

REFERENCIAS

ALVES, L. Apresentação: construindo novos olhares para a presença da inteligência artificial nos cenários educativos. 2023. In.: ALVES, Lynn (Org.). **Inteligência artificial e educação: refletindo sobre os desafios contemporâneos**. – Salvador: EDUFBA; Feira de Santana: UEFS Editora, 2023.

ALVES, L. Notas iniciais sobre inteligência artificial e educação. 2023. In.: ALVES, Lynn (Org.). **Inteligência artificial e educação: refletindo sobre os desafios contemporâneos**. – Salvador: EDUFBA; Feira de Santana: UEFS Editora, 2023.

BENÍTEZ, G. S. El uso de las imágenes en la clase E/LE para el desarrollo de la expresión oral y escrita. **Revista Suplementos marcoELE**, v. 8, p. 1-17, 2009.

BARBOSA, J. L. N.; VIEIRA, J. P.; SANTOS, R. L.; JUNIOR, G. V.; MUNIZ, M. S.; MOURA, R. S. Introdução ao processamento de linguagem natural usando python. **III Escola Regional de Informática do Piauí**. Livro Anais - Artigos e Minicursos, v. 1, n. 1, p. 336-360, jun, 2017.

BRASIL. **Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)**. Brasília: Presidência da República, [2015]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 22 jun. 2022.

BRASIL. **Lei nº 15.249, de 3 de novembro de 2025**. Altera a Lei nº 10.098, de 19 de dezembro de 2000 (Lei da Acessibilidade), e a Lei nº 13.146, de 6 de julho de 2015 (Estatuto da Pessoa com Deficiência), para dispor sobre a instalação de sistemas de comunicação aumentativa e alternativa de baixa tecnologia em espaços públicos e abertos ao público, com vistas à promoção da acessibilidade da pessoa com necessidades complexas de comunicação. Brasília: Presidência da República, 2025.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília, 2018. Disponível em: http://basenacionalcomum.mec.gov.br/images/historico/BNCC_EnsinoMedio_embaixa_site_110518.pdf. Acesso em: 22 nov. 25.

BRASIL. **NOTA TÉCNICA Nº 21 / 2012 / MEC / SECADI /DPEE** Data: 10 de abril de 2012 Assunto: Orientações para descrição de imagem. Disponível em: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=10538-nota-tecnica-21-mecdaisy-pdf&Itemid=30192. Acesso em 03 jul. 2022.

BRASIL. Presidência da República. Secretaria de Direitos Humanos. Secretaria Nacional de Promoção dos Direitos da Pessoa com Deficiência. **Convenção dos direitos das pessoas com deficiência**. 4. ed. Brasília: Secretaria Nacional de Promoção dos Direitos da Pessoa com Deficiência, 2011. 102 p.

BRASIL. Ministério da Saúde. **OMS alerta que 285 milhões de pessoas no mundo têm a visão prejudicada**. Brasília, DF, 2023. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/noticias/2023/fevereiro/oms-alerta-que-285-milhoes-de-pessoas-no-mundo-tem-a-visao-prejudicada>. Acesso em: 28 maio 2026.

CANTERO, F. J. **El arte de no enseñar lengua**. Octaedro: Barcelona, 2019.

CERCHIARI, C. M. **Deficiência visual e ensino/aprendizagem de língua estrangeira**: subsídios para a formação inicial de professores em contexto universitário. Dissertação (Mestrado – Programa de Pós-Graduação em Educação) - Faculdade de Educação da Universidade de São Paulo, 2011.

CHANG, M. H. **O poder da imagem na aprendizagem de língua estrangeira**. 2019. Dissertação de Mestrado. Universidade Católica Portuguesa (Portugal), 2019.

CHIZZOTTI, A. **Pesquisa qualitativa em ciências humanas e sociais**.- Petrópolis, RJ: Vozes, 2006.

CONSEJO DE EUROPA. **Marco de referencia europeo para el aprendizaje, la enseñanza y la evaluación de lenguas**, 2001. Disponível em: https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco. Acesso em: maio de 2026.

DIÁRIO PCD. Brasil tem cerca de 18,6 milhões de pessoas com deficiência, aponta PNAD e IBGE. 2023. Disponível em: <https://diariopcd.com.br/brasil-tem-cerca-de-186-milhoes-de-pessoas-com-deficiencia-aponta-pnad-e-ibge>. Acesso em: 28 maio 2026.

FRANCO, M. M. S.; ALMEIDA FILHO, J. C. P. O conceito de competência comunicativa em retrospectiva e perspectiva. **Revista desempenho**, v. 1, n. 11, 2009.

JAKOBSON, R. **Aspectos da Tradução**. Trad. Izidoro Bilikstein. In: JAKOBSON, R. Linguística e Comunicação. São Paulo: Cultrix, 1995. p. 36-86.

GALASSO, B. Formação Docente para IA Generativa e Acessibilidade. **REVISTA INTERSABERES**, [S. l.], v. 21, p. e26do202, 2026. DOI: 10.22169/revint.v21.e26do202. Disponível em: <https://www.revistasuninter.com/intersaberes/index.php/revista/article/view/2979>. Acesso em: 15 jun. 2026.

GOMES, D. S. Inteligência Artificial: conceitos e aplicações. **Revista Olhar Científico**, v. 1, n. 2, p. 234-246, 2010.

GONSALVES, E. P. **Iniciação à pesquisa científica**. 3. ed. Campinas: Alínea, 2003.

MALTA, G. A pesquisa em audiodescrição realizada em nível de mestrado e doutorado no Brasil nas duas primeiras décadas do século XXI: um estudo quantitativo. *In*: RIBEIRO, E.; ANTUNES, M. A. G (Orgs). **Tradução: história, historiografia, análises textuais e culturais**. Campinas: Pontes, 2026.

MALTA, G.; SCARABELE-ELIS, L. S. Audiodescrição e ensino-aprendizagem de línguas: uma relação mais que benéfica, necessária. *In*.: MALTA, Gleiton (org.). **Ensino de tradução e de línguas: reflexões e propostas didáticas**. Campinas: Pontes, 2023.

MEDRADO, B. P. **Deficiência visual e ensino de línguas estrangeiras: políticas, formação e ações inclusivas**. Campinas: Pontes Editora, 2014.

MIANES, F. L. A audiodescrição nos processos de ensino e aprendizagem na Educação Básica. **Revista Linguagem em Foco**, v. 15, n. 2, p. 30-46, 2023.

MOTTA, L. M. **A audiodescrição na escola: abrindo caminhos para leitura de mundo**. Pontes Editora. Campinas, 2016.

MOTTA, L. M; ROMEU FILHO, P. **Audiodescrição: transformando imagens em palavras**. São Paulo: Secretaria dos direitos da pessoa com deficiência do Estado de São Paulo, p. 7, 2010.

NAVES, S. B.; MAUCH, C.; ALVES, S.; ARAÚJO, V. L. S. **Guia para produções audiovisuais acessíveis**. Brasília: Secretaria do Audiovisual do Ministério da cultura, 2016.

NORD, C. El funcionalismo en la enseñanza de traducción. *Mutatis Mutandis*. **Revista latinoamericana de traducción**, v. 2, n. 2, p. 209-243, 2009.

NORD, C. **Texto base-texto meta: un modelo funcional de análisis pretraslativo**. Tradução e adaptação de Cristiane Nord. Castelló de la Plana: Publicacions de la Universitat Jaume I, 2012.

PEREIRA, A. C. P. O uso da inteligência artificial na educação: possibilidades e limitações. **Revista de Inovação, Tecnologia e Educação**, São Paulo, v. 5, n. 1, 2018.

PEREIRA, S. L. **Processamento de linguagem natural**, 2019. Disponível em: <https://www.ime.usp.br/~slago/IA-pln.pdf>. Acesso em: 2 de jun. 2024.

PERDIGÃO, L. T. *et al*. Inteligência artificial para audiodescrição de imagens: uma análise da pessoa com deficiência visual. *In*.: **Anais do VIII Congresso sobre Tecnologias na Educação**. SBC, 2023. p. 182-191.

PERIS, Ernesto Martín et al. **Gente única**. 1. ed. Barcelona: Difusión, 2018.

PONTES, V. O; PEREIRA, L. O. **A tradução a partir do modelo funcionalista de Christiane Nord: perspectivas para o ensino de línguas estrangeiras**. *Revista Tradterm*, v. 28, p.338-363, 2016.

REISS, K.; VERMEER, H. J. Fundamentos para una teoría funcional de la traducción. Tradução de Sandra García Reina e Celia Martín de León. Madrid: Ediciones Akal, 1996. 206p.

SANTAELLA, L. **A inteligência artificial é inteligente?** São Paulo: Edições 70, 2023.

SILVA, J. L.; FONSECA, R. S.; SILVA, T. C.; SILVA, V. B.; RIBEIRO, E. N. A inteligência artificial como ferramenta audiodescritora para estudantes com deficiência visual. **Anais do V CINTEDI**. Campina Grande: Realize Editora, 2024. Disponível em: <https://editorarealize.com.br/artigo/visualizar/108150>. Acesso em: 08/02/2025.

SOUSA, I. V. Audiodescrição: o que é? Como se faz? **Revista EDaPECI**, v. 17, n. 3, p. 34-45, 2017.

SCARABELE-ELIS, L. **A audiodescrição realizada por ferramentas de inteligência artificial: um estudo a partir de imagens de um livro didático de espanhol como língua estrangeira**. 2025. 241 f., il. Dissertação (Mestrado em Estudos da Tradução) – Universidade de Brasília, Brasília, 2025.

SCARABELE-ELIS, L. **A audiodescrição como ferramenta pedagógica no ensino de espanhol como língua estrangeira**. 2022. Trabalho de Conclusão de Curso (Graduação em Letras Espanhol e Literaturas Espanhola e Hispano-Americana) – Departamento de Línguas Estrangeiras e Tradução, Instituto de Letras, Universidade de Brasília, Brasília, 2022.