



Práctica de pares mínimos en inglés con Teachable Machine: impacto en el reconocimiento auditivo y la producción oral en primaria

Jesica Alejandra Cortés Castro

Diego Alejandro Cortes Londoño

María Alejandra Nieto León

Luz Estela Valencia Valencia

Trabajo de grado presentado para optar al título de Magíster en Educación y Transformación Digital

Asesor: Carlos Betancourt Correa, Doctor (PhD) en Ciencias de la Educación

Asesores de recursos académicos: Juan Pablo Charry Osorio (asesor bibliográfico), Claudia Marcela Cerón Rubio (asesora Centro de Escritura) y Elvia Lucía Sánchez García (asesora de integridad académica)

Universidad de Manizales
Facultad de Ciencias e Ingeniería
Maestría en Educación y Transformación Digital - Virtual
Manizales, Caldas, Colombia

2025

Cita	(Cortés Castro et al, 2025)
Referencia	Cortés Castro, J. A.; Cortés Londoño, D. A.; Nieto León, M. A.; Valencia Valencia L. E. (2025). Práctica de Pares Mínimos en Inglés con Teachable Machine: Impacto en Reconocimiento Auditivo y Producción Oral en Primaria. [Tesis de maestría]. Universidad de Manizales. RIDUM: Repositorio Institucional Universidad de Manizales.
Estilo APA 7 (2020)	



Maestría en Educación y Transformación Digital - Virtual, IV

Declaración de inteligencia artificial: Los autores de este trabajo de grado declaran que se emplearon herramientas de inteligencia artificial (ChatGPT, NotebookLM, SciSpace, y Gemini) de manera ética y responsable, conforme al Acuerdo UManizales 002 (julio 26 de 2023). Estas herramientas apoyaron la redacción preliminar, la organización de ideas, la búsqueda bibliográfica, la verificación de estilo y la codificación de análisis en Python, pero en ningún caso sustituyeron el análisis crítico, la interpretación de resultados ni la originalidad del trabajo. Para la verificación es importante informar que se utilizó el detector de IA de SciSpace, herramienta especializada que se consideró pertinente porque uno de los investigadores contaba con membresía activa. No obstante, esta aplicación dejó de aportar información rigurosa al identificar como “generado por IA” fragmentos elaborados directamente por los autores y, en etapas posteriores, se limitó a producir reportes poco precisos. Ninguna otra herramienta ofreció una opción viable y aplicable para “humanizar” el texto, procedimiento que no se considera un uso responsable de la IA si lo hace la IA misma. El presente es un trabajo original asistido por Inteligencia Artificial, lo suficientemente complejo como para que ninguna IA lo hubiera podido generar sin la mediación activa de los investigadores. Toda salida generada por IA fue contrastada con las fuentes primarias, discutida en grupo y registrada en la Bitácora de IA (Anexo 11), que garantiza la trazabilidad del proceso.

Biblioteca y Centro de Recursos: <https://biblioteca.umanizales.edu.co/>

Repositorio Institucional: <http://ridum.umanizales.edu.co/>

Universidad de Manizales: www.umanizales.edu.co

Revistas: <http://revistasum.umanizales.edu.co/>

Fondo Editorial: <https://editorialum.umanizales.edu.co/>

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Manizales ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Dedicatoria

A todos aquellos que tienen sueños en su vida por cumplir, no se rindan, nosotros estamos culminando el proceso con el fin de alcanzar el sueño que algún día tuvimos. Sigamos adelante y no desfallezcan en el proceso, y recuerden siempre **TODO LO QUE SUEÑEN LO PODRÁN ALCANZAR.**

Agradecimientos

Queremos expresar nuestro más profundo agradecimiento a todas las personas e instituciones involucradas en el desarrollo de este proyecto de investigación.

En primer lugar, a los niños y niñas participantes, quienes con entusiasmo, disposición y alegría aportaron su tiempo y esfuerzo para llevar a cabo las actividades del proyecto. Su participación activa constituyó el motor principal de este trabajo. Nuestro agradecimiento también a la Institución Educativa INEM José Celestino Mutis, por abrir sus puertas y permitirnos implementar esta propuesta en su comunidad educativa, contribuyendo así al avance de la investigación y a la innovación pedagógica.

Agradecemos también a las familias, por su apoyo constante, la confianza depositada y la disposición para facilitar la participación de sus hijos en cada etapa de la investigación. Al profesor Carlos Betancourt quien brindó orientación, comentarios valiosos y un acompañamiento cercano durante el desarrollo de este estudio. Asimismo, expresamos nuestra gratitud al docente evaluador externo, quien dedicó su tiempo a la calificación de los audios y ofreció una perspectiva objetiva que fortaleció la validez de los resultados obtenidos.

Finalmente, reconocemos a Google y sus herramientas gratuitas, entre ellas Teachable Machine, que facilitaron la implementación de la propuesta didáctica, demostrando que la tecnología, cuando se emplea con sentido pedagógico, puede convertirse en una aliada poderosa para la transformación educativa.

Tabla de contenido

Resumen	12
Abstract	13
Introducción	14
1 Planteamiento del problema	16
1.1 Antecedentes	17
1.1.1 Dificultades de pronunciación en hispanohablantes	17
1.1.2. Sistemas CAPT y retroalimentación automatizada.....	17
1.1.3. Evidencia en contextos escolares y universitarios	18
1.1.4. Revisión de literatura y hallazgos globales.....	18
1.1.5. Teachable Machine de Google: potencial y limitaciones	19
1.1.6. Vacío identificado y aporte de la investigación	19
2 Justificación.....	20
3 Objetivos	22
3.1 Objetivo general	22
3.2 Objetivos específicos.....	22
4 Hipótesis.....	23
4.1 Hipótesis de trabajo	23
4.1.1 Hipótesis nula.....	24
4.1.1.1 Hipótesis alterna.	24
4.1.1.1.1 Variables.....	24
5 Marco teórico	26
5.1 Fundamentos pedagógicos: construccionismo y aprendizaje activo.....	26
5.1.1. Construccionismo y aprendizaje activo	26
5.1.2. Potencial de la IA en la motivación y el rendimiento	27

5.1.3. Rol docente y andamiaje en entornos con IA	29
5.2. Enfoque lingüístico y categorías de la investigación	30
5.2.1. Categorías centrales de la investigación	30
5.2.1.1 Reconocimiento auditivo.....	30
5.2.1.2 Producción oral.	30
5.2.2. Teoría fonética, pares mínimos y conciencia fonológica.....	30
5.2.2.1 Contrastes vocálicos críticos.	31
5.2.2.2 Contrastes consonánticos críticos.....	31
5.2.3. El “Noticing” en la adquisición de L2	32
5.2.4. Percepción del estudiante en procesos mediados por IA.....	33
5.3 Integración tecnológica y fundamentación de la propuesta	34
5.3.1 Machine Learning y Teachable Machine.....	34
5.3.2. Comparación de enfoques: tradicional vs. mediado por IA.....	34
5.3.2.3. Accesibilidad e inclusión.	35
5.3.3. Competencias digitales y pensamiento computacional.....	36
5.3.3.1. Alfabetización digital.	36
5.3.3.2. Recolección y análisis de datos.....	36
5.3.4. Limitaciones y consideraciones éticas del uso de IA.....	36
5.3.4.1. Consideraciones éticas y privacidad de datos.	37
5.3.4.2. Sesgos en los datos.....	37
5.3.4.3. Comprensión limitada del contexto y la inteligencia emocional.	37
5.3.4.4. Brecha digital.	37
5.3.4.5. Limitaciones técnicas de ASR.....	37
5.4. Síntesis y pertinencia para la investigación.....	38
6 Metodología	39

6.1 Enfoque de investigación	39
6.2 Diseño de investigación	39
6.3 Contexto y participantes.....	39
6.3.1 Criterios de inclusión	40
6.3.2 Criterios de exclusión	40
6.4 Variables y categorías de estudio	41
6.4.1 Variable independiente.	41
6.4.1.1 Propuesta didáctica con GTM.	41
6.4.2 Variables dependientes	41
6.4.2.1 Reconocimiento auditivo de pares mínimos en inglés	41
6.4.2.2. Producción oral de pares mínimos en inglés.....	41
6.4.2.3 Percepción frente a la estrategia.....	42
6.5 Instrumentos de recolección de datos.....	42
6.5.1 Prueba de reconocimiento auditivo y producción oral de pares mínimos.	42
6.5.2 Encuesta de percepción.	43
6.5.3 Observación docente.	43
6.5.4 Registros de desempeño en GTM	43
6.5.5 Validez y confiabilidad	43
6.6 Procedimiento.....	44
6.6.1. Fase 1: Pretest	44
6.6.1.1 Reconocimiento auditivo.....	44
6.6.1.2 Producción oral.	44
6.6.2. Fase 2: Intervención didáctica	44
6.6.3. Fase 3: Postest.....	45
6.6.4. Fase 4: Análisis y sistematización	45

6.7	Consideraciones éticas	46
6.8	Integridad académica y uso de IA	48
7.	Resultados y análisis	50
7.1	Comparabilidad inicial de los grupos.....	50
7.2	Verificación de supuestos estadísticos y fiabilidad intercalificador	50
7.3	Análisis de los cambios intragrupo	51
7.3.1	Sección A: Reconocimiento auditivo.....	51
7.3.2	Sección B: Producción oral.....	52
7.4	Análisis de comparación entre grupos	53
7.5	Análisis de heterogeneidad y desempeño por ítems.....	53
7.6	Encuesta de percepción del GE.....	56
7.6.1	Dimensión 1: Motivación y disfrute	56
7.6.2	Dimensión 2: Autoeficacia y confianza.....	57
7.6.3	Dimensión 3: Usabilidad y utilidad percibida	57
7.6.4	Dimensión 4: Experiencia constructora.....	57
7.7.	Observaciones sistemáticas del docente investigador	58
8	Discusión.....	60
8.1	Efectividad de la intervención y la ausencia de diferencias significativas.....	60
8.2	Construccionismo: motivación y aprendizaje activo.....	61
8.3	Resultados por categorías fonológicas	62
8.4	Dimensión afectivo-motivacional y comparación con la literatura.....	62
9	Conclusiones	64
10	Recomendaciones.....	67
10.1	Recomendaciones para la práctica pedagógica	67
10.1.1	Brevidad de la intervención.....	67

10.1.2 Dependencia tecnológica y desafíos logísticos.....	67
10.1.3 Precisión de los modelos creados por estudiantes	68
10.1.4 Adoptar un modelo pedagógico híbrido y flexible	68
10.1.5 Diseñar una transición del trabajo colaborativo al individual	68
10.2 Recomendaciones para futuras investigaciones	69
10.2.1 Estudios longitudinales sobre el impacto acumulativo	69
10.2.2 Diseños experimentales con aleatorización	69
10.2.3 Estudios comparativos de herramientas de IA.....	69
10.2.4 Investigación sobre variables afectivas y metacognitivas	69
Referencias	71
Anexos.....	76

Lista de tablas

Tabla 1 Resultados de la Sección A (Reconocimiento Auditivo)	51
Tabla 2 Resultados de la Sección B (Producción oral)	52
Tabla 3 Resultados por categoría fonológica en Sección A.....	54
Tabla 4 Resultados por categoría fonológica en Sección B	55
Tabla 5 Resultados de la encuesta de percepción en GE	58
Tabla 6 Bitácora de uso de IA	103

Lista de figuras

Figura 1 Diagrama Metodológico de la Investigación.....	48
Figura 2 Distribución de ganancias por grupo en la Sección A.....	52
Figura 3 Distribución de ganancias por grupo en Sección B.....	53
Figura 4 Mejora porcentual por categoría fonológica en Sección A.....	55
Figura 5 Mejora porcentual por categoría fonológica en Sección B.....	56
Figura 6 Indicador de frustración por sesión en el GE.....	59

Siglas, acrónimos y abreviaturas

ABP	Aprendizaje Basado en Proyectos
ASR	Reconocimiento Automático del Habla
CAPT	Sistema de Entrenamiento de Pronunciación Asistido por Computadora
EFL	Inglés como Lengua Extranjera
GE	Grupo Experimental
GC	Grupo Control
IA	Inteligencia Artificial
L1	Lengua Materna
L2	Segunda Lengua
MCER	Marco Común Europeo de Referencia para las Lenguas
ML	Machine Learning
PNB	Programa Nacional de Bilingüismo
TIC	Tecnologías de la Información y la Comunicación
GTM	Google Teachable Machine
TTS	Síntesis de Texto a Voz

Resumen

La enseñanza de la pronunciación del inglés en Colombia enfrenta obstáculos notables, particularmente con los pares mínimos que resultan complejos para los hispanohablantes. Esta investigación buscó evaluar el impacto de una propuesta didáctica constructivista, basada en la creación de modelos de clasificación de audio con Google Teachable Machine (GTM), sobre el reconocimiento auditivo y la producción oral en estudiantes de primaria. Se implementó un estudio cuasiexperimental de pretest y posttest con un grupo control (n=21) que recibió instrucción tradicional y un grupo experimental (n=22) que trabajó activamente con la herramienta de IA. Los resultados cuantitativos mostraron que ambos grupos obtuvieron mejoras estadísticamente significativas, sobre todo en la producción oral, sin que existieran diferencias relevantes entre ellos. Sin embargo, los hallazgos cualitativos fueron contundentes: la intervención con GTM generó niveles excepcionales de motivación, participación y percepción positiva del aprendizaje. Se concluye que, si bien la efectividad de la propuesta es comparable a la de un método tradicional bien ejecutado en el corto plazo, su principal valor radica en su capacidad para transformar la experiencia educativa, fomentando un aprendizaje activo, autónomo y altamente motivador.

Palabras clave: Inteligencia Artificial, constructivismo, aprendizaje activo, Teachable Machine, pares mínimos, pronunciación en inglés, educación primaria.

Abstract

Teaching English pronunciation in Colombia faces significant obstacles, particularly with minimal pairs that are challenging for Spanish speakers. This research aimed to evaluate the impact of a constructionist didactic proposal, based on creating audio classification models with Google Teachable Machine (GTM), on the auditory recognition and oral production of primary school students. A quasi-experimental pre-test/post-test study was implemented with a control group (n=21) that received traditional instruction and an experimental group (n=22) that actively worked with the AI tool. Quantitative results showed that both groups achieved statistically significant improvements, especially in oral production, with no significant differences between them. However, the qualitative results were conclusively impactful on the experimental group: the GTM intervention generated exceptional levels of motivation, engagement, and positive learning perceptions. The study concludes that while the proposal's effectiveness is comparable to a well-executed traditional method in the short term, its main value lies in its capacity to transform the educational experience by fostering active, autonomous, and highly motivating learning.

Keywords: Artificial Intelligence, constructionism, active learning, Teachable Machine, minimal pairs, English pronunciation, primary education.

Introducción

En el panorama educativo de Colombia, la enseñanza del inglés como lengua extranjera (EFL, *English as a Foreign Language*) presenta desafíos persistentes que limitan el desarrollo de una competencia comunicativa real en la educación básica. Esta investigación nace de la necesidad de abordar una de las dificultades más arraigadas para los estudiantes hispanohablantes: la correcta pronunciación y el reconocimiento auditivo de sonidos en inglés que no existen en su lengua materna; siendo este uno de los mayores desafíos donde la educación tradicional muestra falencias a lo largo del ciclo educativo en la enseñanza de inglés como lengua extranjera (EFL).

El presente estudio se centra en el contexto específico de los estudiantes de quinto grado de la Institución Educativa INEM José Celestino Mutis. Al igual que en muchas escuelas públicas, el aprendizaje del inglés se caracteriza por una exposición limitada al idioma y por una dependencia de modelos pedagógicos que no siempre fomentan la interactividad o la retroalimentación inmediata, elementos cruciales para superar las barreras fonológicas entre el español y el inglés. Frente a un escenario nacional donde los avances en bilingüismo han sido modestos, se vuelve fundamental la búsqueda de nuevas alternativas metodológicas.

Para responder a esta problemática, este trabajo se propone explorar el potencial de una propuesta didáctica que integra la inteligencia artificial (IA) desde un enfoque innovador. Así, se introduce GTM, una herramienta de aprendizaje automático (ML, *Machine Learning*) que sitúa al estudiante en un rol activo, no como un simple consumidor de tecnología, sino como un creador de sus propios modelos de clasificación de audio. A través de este enfoque, que se fundamenta en la teoría del construccionismo, se busca ir más allá de la práctica fonética, con la intención de potenciar también la motivación intrínseca y el pensamiento computacional.

Por consiguiente, el objetivo general de esta investigación es evaluar el efecto de esta propuesta didáctica sobre el reconocimiento auditivo y la producción oral de pares mínimos en inglés, dentro de un diseño cuasiexperimental con pretest y posttest, contrastando los resultados de un GE con los de un grupo de control que siguió una instrucción tradicional.

La justificación de este estudio es triple. Desde una perspectiva teórica, busca aportar evidencia a un campo emergente que combina la IA y los enfoques construccionistas en la educación primaria. En el plano práctico, ofrece una estrategia didáctica tangible y replicable, diseñada para enriquecer la labor docente en contextos con recursos limitados. Finalmente, en su

dimensión social, esta investigación se alinea con la aspiración nacional de fortalecer el bilingüismo, al tiempo que promueve una alfabetización digital crítica y necesaria para los ciudadanos del siglo XXI.

1 Planteamiento del problema

Esta investigación se lleva a cabo en la Institución Educativa INEM José Celestino Mutis, una escuela pública urbana ubicada en la Comuna 9 de la ciudad de Armenia, Quindío. Con más de 50 años de trayectoria, la institución es un referente académico regional que ofrece educación desde preescolar hasta media técnica y académica en jornada única. De manera crucial para este estudio, cuenta con una infraestructura tecnológica adecuada, incluyendo salas de informática, que facilitan la implementación de estrategias didácticas mediadas por herramientas digitales. El estudio se centra específicamente en estudiantes de quinto grado de primaria, quienes cursan EFL bajo un modelo de enseñanza predominantemente tradicional.

A nivel nacional, la enseñanza del inglés en Colombia enfrenta desafíos estructurales significativos. Políticas educativas como el Programa Nacional de Bilingüismo (PNB) no han logrado alcanzar sus metas, lo que ha resultado en un estancamiento generalizado en el dominio del idioma. Benavides (2021) reporta que, mientras se esperaba que el 40 % de los estudiantes de grado 11 alcanzaran el nivel B1 para 2014, la cifra real fue solo del 4 %, aumentando apenas al 7 % en 2017. Este panorama se refleja también en la prueba Saber Pro, donde la mejora anual promedio en los niveles B1 y B+ fue de solo 0.1 % en una década (Benavides, 2021). La problemática se agrava por la limitada exposición al inglés fuera del aula, con horarios reducidos (1–2 horas semanales) y escasas oportunidades de práctica en contextos cotidianos (Guerrero-Nieto, 2024).

En este marco, los estudiantes hispanohablantes enfrentan dificultades específicas en el reconocimiento auditivo y la producción oral debido a diferencias fonológicas entre español e inglés, especialmente en el sistema vocálico y en los patrones prosódicos (Silva Valencia, 2022; Acosta Acosta, 2023; Uribe-Enciso et al., 2019). Los métodos de enseñanza tradicionales, basados en repetición y ejercicios escritos, no ofrecen la retroalimentación inmediata ni la interactividad necesarias para superar estas barreras (Tejedor-García, 2020; Neri, 2007).

Ante esta situación, la presente investigación plantea el uso de GTM, una herramienta de ML accesible y de fácil uso, que permite a los estudiantes diseñar y programar modelos de clasificación de audio. Al promover la retroalimentación inmediata y el aprendizaje activo, GTM se alinea con enfoques pedagógicos constructivistas y ofrece una oportunidad de innovación en la enseñanza del inglés en primaria en Colombia.

1.1 Antecedentes

1.1.1 Dificultades de pronunciación en hispanohablantes

La integración de la IA en la educación ha cobrado una relevancia creciente, particularmente en la enseñanza de EFL. Para los hispanohablantes, la pronunciación supone un reto debido a diferencias fonológicas entre el español y el inglés, incluyendo fonemas vocálicos y consonánticos, diptongos, triptongos y patrones de acentuación (Silva Valencia, 2022). Estos desafíos en la enseñanza requieren enfoques didácticos y estrategias metodológicas orientadas a la personalización de la retroalimentación significativa y puntual que posibilite una mejora por parte de los aprendices. En este contexto, la IA, mediante tecnologías como el Reconocimiento Automático del Habla (ASR) y la Síntesis de Texto a Voz (TTS), se ha consolidado como un recurso prometedor para apoyar el entrenamiento de la pronunciación (Tejedor-García et al., 2020).

1.1.2. Sistemas CAPT y retroalimentación automatizada

Diversas investigaciones han mostrado que los Sistemas de Entrenamiento de Pronunciación Asistidos por Computadora (CAPT) que incorporan ASR y TTS alcanzan resultados significativos. Por ejemplo, Tejedor-García et al. (2020) reportaron que el sistema de reconocimiento de voz de Google alcanzó un nivel de precisión del 95 % en inglés, cercano al umbral humano. En su estudio experimental con hablantes nativos de español, nivel B1-B2 de inglés según el MCER, el grupo que utilizó CAPT durante cuatro semanas incrementó en promedio 1.74 puntos su desempeño, lo que representó una mejora relativa del 205 %, frente al 82 % del grupo con instrucción tradicional.

En otros prototipos como *English Vowels*, *Japañol* y *COP*, los estudiantes incrementaron sus tasas de éxito en discriminación y producción de sonidos, con diferencias estadísticamente significativas frente a quienes recibieron instrucción tradicional. Se observó además que los aprendices con menor dominio inicial aprovecharon más intensamente los modelos TTS, logrando avances mayores en reconocimiento y producción (Tejedor-García et al., 2020).

1.1.3. Evidencia en contextos escolares y universitarios

La eficacia de estos sistemas también ha sido confirmada en diferentes escenarios educativos. El sistema PLASER, implementado con 900 estudiantes cantonenses, mostró que el 73 % de los usuarios mejoró su precisión de pronunciación con un incremento absoluto promedio del 4.53 %, estadísticamente significativo al 99 % de confianza (Mak et al., 2003).

En Europa, Neri et al. (2008) hallaron que niños de 11 años mejoraron la calidad segmental de 11 fonemas problemáticos, con un incremento de la puntuación media de 4.49 a 6.67 en el grupo que recibió retroalimentación automática. De manera similar, Pennington y Rogerson-Revell (2019) demostraron que incluso entrenamientos simples con ASR pueden generar mejoras comparables a las de contextos presenciales, confirmando la efectividad de la retroalimentación automatizada.

En escenarios escolares latinoamericanos, López-Minotta et al. (2025) evidenciaron que estudiantes de quinto grado mejoraron significativamente en pronunciación, de 6.40 a 13.20 puntos, y en expresión oral, de 3.28 a 7.12, tras 16 semanas de intervención. Asimismo, en educación superior, Senowarsito y Ardini (2023) mostraron que el uso de la aplicación ELSA Speak con 32 universitarios incrementó la puntuación promedio de pronunciación de 60.06 a 85.63 en nueve sesiones, aunque con limitaciones en el entrenamiento de rasgos suprasegmentales.

1.1.4. Revisión de literatura y hallazgos globales

Las revisiones sistemáticas confirman que las herramientas de IA generan mejoras en aspectos segmentales y suprasegmentales de la pronunciación (Aryanti y Santosa, 2024). Además, técnicas tradicionales como los pares mínimos potencian su impacto. Acosta Acosta (2023), en un estudio con 33 universitarios, reportó un incremento de 6.37 a 7.43 puntos tras diez sesiones, evidenciando el valor de esta técnica para la discriminación auditiva y la producción oral.

A mayor escala, Forero-Corba y Negre Bennasar (2024) identificaron que el 74.6 % de los estudios aplicaron IA en primaria y secundaria, encontrando 33 técnicas de ML aplicadas en educación, algunas con una capacidad predictiva superior al 96 %, lo que demuestra el potencial de estas tecnologías para personalizar procesos de aprendizaje.

1.1.5. Teachable Machine de Google: potencial y limitaciones

En este marco, GTM emerge como una alternativa novedosa que trasciende la práctica pasiva del estudiante y lo convierte en programador de sus propios modelos de clasificación. GTM introduce un componente constructorista en el que los aprendices participan en la recolección de datos, el etiquetado de audios y el entrenamiento de modelos (Carney et al., 2020; Sayginer y Tekdal, 2023). Este giro metodológico resulta relevante porque no solo permite la práctica fonética, sino que involucra al estudiante en principios de ML, fortaleciendo la alfabetización digital y el pensamiento computacional (Dwivedi et al., 2024).

Estudios recientes han documentado la viabilidad de GTM en educación básica y media, destacando su facilidad de uso, accesibilidad y bajo costo (Pujari et al., 2022). También se ha evidenciado que favorece el aprendizaje activo y la motivación estudiantil al integrar tareas de programación y experimentación con datos reales (Dwivedi et al., 2024). No obstante, persisten limitaciones técnicas, como el reconocimiento en ambientes con ruido y la necesidad de grandes volúmenes de datos para mejorar la precisión (Pujari et al., 2022).

1.1.6. Vacío identificado y aporte de la investigación

Sayginer y Tekdal (2023) subrayan que, pese al potencial educativo de GTM, su uso sigue siendo limitado y la investigación aún se encuentra en etapas iniciales. Además, su aplicación se ha concentrado en la enseñanza de conceptos de ML en contextos K-12, principalmente relacionados con alfabetización digital y ciencias.

En síntesis, la evidencia demuestra que los sistemas CAPT y las aplicaciones de IA producen mejoras significativas en la pronunciación en diversos contextos educativos. Sin embargo, la mayoría de los estudios se han desarrollado con estudiantes mayores o en entornos universitarios, y el uso de enfoques constructoristas como GTM en la enseñanza de la pronunciación mediante pares mínimos en primaria es prácticamente inexistente. Este vacío justifica la pertinencia de explorar la efectividad de una estrategia didáctica basada en GTM para fortalecer el reconocimiento auditivo y la producción oral en estudiantes de quinto grado, aportando evidencia original y necesaria en la enseñanza de lenguas mediada por tecnologías emergentes.

2 Justificación

La enseñanza del inglés en Colombia continúa enfrentando limitaciones estructurales que afectan directamente el desarrollo de competencias comunicativas, en especial en contextos escolares públicos donde la exposición al idioma es mínima. Esta realidad se refleja en los bajos resultados nacionales de dominio de inglés, evidenciados por los reportes del PNB y las pruebas estandarizadas (Benavides, 2021; Guerrero-Nieto, 2024).

Más allá de señalar estas dificultades, esta investigación busca aportar una respuesta innovadora y contextualizada al problema del reconocimiento auditivo y la producción oral en estudiantes de quinto grado del INEM José Celestino Mutis, integrando además el análisis de la percepción de los estudiantes frente a la estrategia como un aporte cualitativo complementario. El uso de GTM introduce un componente constructor al permitir que los estudiantes no solo practiquen pares mínimos, sino que también programen, experimenten y reciban retroalimentación inmediata, favoreciendo el aprendizaje activo.

La pertinencia del estudio se justifica en tres dimensiones:

1. **Teórica:** Aporta a la consolidación de la literatura sobre integración de IA en la enseñanza de lenguas, un campo en desarrollo donde aún no se ha explorado suficientemente el uso de GTM para el entrenamiento fonético en educación primaria (Sayginer y Tekdal, 2023). De este modo, se conecta la teoría del constructorismo con la práctica fonética mediada por tecnologías emergentes.

2. **Práctica:** Propone una estrategia didáctica replicable y viable en instituciones públicas con infraestructura tecnológica básica. Al aprovechar GTM, se plantea un modelo de enseñanza que supera la repetición mecánica y ofrece a los docentes una alternativa metodológica innovadora y sostenible para trabajar las dificultades fonológicas de los estudiantes. Asimismo, el diseño contempla la exploración de la percepción estudiantil como insumo para valorar la motivación, la satisfacción y la viabilidad pedagógica de la propuesta.

3. **Social:** Responde a las políticas nacionales de bilingüismo y al llamado de integrar TIC en la educación. Además, al enfocarse en la pronunciación como competencia clave para la

comunicación, contribuye a fortalecer la equidad educativa y la alfabetización digital en contextos escolares, potenciando oportunidades futuras para los estudiantes.

En síntesis, este estudio no solo atiende un problema local y nacional en la enseñanza del inglés, sino que también genera conocimiento novedoso en la intersección entre fonética, pedagogía construccionista e IA aplicada, con impacto académico, práctico y social.

3 Objetivos

3.1 Objetivo general

Evaluar el efecto de una propuesta didáctica de cuatro sesiones, centrada en el aprendizaje activo con GTM, sobre el reconocimiento auditivo y la producción oral de pares mínimos en inglés, en un GE de estudiantes de quinto grado del INEM José Celestino Mutis, midiendo el cambio en sus puntuaciones promedio mediante pruebas pretest y posttest, e incorporando análisis estadísticos de significancia, tamaños del efecto e intervalos de confianza, y comparando estos resultados con los de un grupo de control que recibió instrucción fonética tradicional.

3.2 Objetivos específicos

- Diseñar e implementar una propuesta didáctica de cuatro sesiones de 80 minutos cada una, basada en el construccionismo y el aprendizaje activo con GTM, para mejorar el reconocimiento auditivo y la producción oral de contrastes fonológicos críticos en el GE.
- Cuantificar, mediante la aplicación de un pretest y un posttest, la mejora en las puntuaciones promedio del GE en el reconocimiento auditivo y la producción oral de pares mínimos, analizando la significancia estadística del cambio con pruebas t para muestras relacionadas, reportando valores de p , tamaños del efecto y estadísticos descriptivos.
- Comparar, utilizando pruebas t para muestras independientes, la efectividad relativa de la propuesta didáctica con GTM frente a la instrucción fonética tradicional, contrastando las puntuaciones promedio del posttest entre el GE y el grupo de control, e interpretando los tamaños del efecto para determinar la magnitud de las diferencias.
- Analizar la percepción de los estudiantes del GE sobre la propuesta didáctica mediada por GTM, identificando niveles de motivación, satisfacción y utilidad pedagógica a partir de encuestas con escala Likert y del registro cualitativo de observación docente.

4 Hipótesis

Dado que esta investigación tiene un alcance explicativo y busca medir el efecto de una intervención, es pertinente formular hipótesis que anticipen los resultados esperados. Se espera que la implementación de una propuesta didáctica innovadora, basada en aprendizaje activo con GTM, genere una mejora estadísticamente significativa en el aprendizaje fonético de los estudiantes y una percepción positiva de la estrategia, en comparación con los métodos tradicionales.

4.1 Hipótesis de trabajo

La implementación de una propuesta didáctica de cuatro sesiones, centrada en el aprendizaje activo mediante la programación de modelos de clasificación de audio en GTM, generará:

- Una mejora estadísticamente significativa en el reconocimiento auditivo y la producción oral de pares mínimos en inglés en el GE, en comparación con el grupo de control que recibe instrucción fonética tradicional, evidenciada a través de diferencias en medias, pruebas t de significancia, tamaños del efecto e intervalos de confianza.
- Una percepción más favorable por parte de los estudiantes del GE frente a la estrategia mediada por GTM, reflejada en niveles más altos de motivación, satisfacción y utilidad pedagógica medidos con encuestas tipo Likert (media, DE) y registros de observación docente.

Esta suposición se fundamenta en la teoría del construccionismo y en evidencia previa que demuestra que los sistemas CAPT mejoran de manera significativa tanto la discriminación como la producción de sonidos en comparación con la instrucción tradicional. La retroalimentación inmediata que ofrece la herramienta GTM facilita la autocorrección y promueve un procesamiento más profundo del material, lo que debería traducirse en un mayor rendimiento y una experiencia de aprendizaje más positiva.

4.1.1 Hipótesis nula

No existirá una diferencia estadísticamente significativa en la mejora del reconocimiento auditivo ni en la producción oral de pares mínimos en inglés entre el GE, que utiliza la propuesta didáctica con GTM, y el grupo de control, que recibe instrucción fonética tradicional. Asimismo, no se observarán diferencias relevantes en la percepción de los estudiantes respecto a motivación, satisfacción y utilidad pedagógica.

4.1.1.1 Hipótesis alterna. La implementación de una propuesta didáctica con GTM generará una mejora estadísticamente inferior en el reconocimiento auditivo y la producción oral de pares mínimos en inglés en el GE, en comparación con el grupo de control, y no favorecerá significativamente la percepción de los estudiantes sobre la experiencia de aprendizaje.

4.1.1.1.1 Variables.

- Variable independiente:

1. Propuesta didáctica con GTM. Corresponde a la intervención pedagógica de cuatro sesiones centrada en el aprendizaje activo y el construccionismo, que utiliza la programación de modelos de clasificación de audio en GTM para la práctica fonética.

- Variables dependientes:

1. Reconocimiento auditivo de pares mínimos en inglés: Capacidad de los estudiantes para discriminar auditivamente contrastes fonológicos críticos como /i:-ɪ/, /u:-ʊ/, /p-b/, /t-d/, /k-g/, /f-v/, /s-z/ y /θ-ð/. Se medirá a través de los puntajes obtenidos en la sección A de la prueba de reconocimiento auditivo y producción oral.

2. Producción oral de pares mínimos en inglés: Nivel de precisión articuladora e inteligibilidad alcanzado al pronunciar los pares mínimos. Se evaluará mediante la rúbrica de la sección B del mismo instrumento.

3. Percepción estudiantil de la estrategia didáctica: Opinión de los estudiantes del GE sobre motivación, satisfacción y utilidad pedagógica de la propuesta con GTM. Se evaluará a través de encuestas con escala Likert y observación docente.

5 Marco teórico

El presente marco teórico establece las bases conceptuales y empíricas que sustentan la investigación sobre el uso del ML en la mejora de la pronunciación del inglés como segunda lengua (L2) en estudiantes de primaria. La problemática de la enseñanza del inglés en el contexto colombiano, caracterizada por la escasez de exposición al idioma, es un factor crítico en el aprendizaje. La insuficiencia de la intensidad horaria y de las oportunidades para la práctica en contextos informales dificulta el desarrollo de competencias comunicativas, especialmente en la producción oral y la escucha (Guerrero-Nieto, 2024; Benavides, 2021). Esta realidad se refleja en las aulas de primaria del INEM José Celestino Mutis, donde los estudiantes hispanohablantes cursan inglés bajo un modelo de enseñanza tradicional con un tiempo limitado de exposición semanal. Esta restricción de tiempo, junto con la falta de interactividad en los métodos convencionales, agrava las dificultades específicas que enfrentan al distinguir y producir pares de sonidos similares. Por lo tanto, cualquier estrategia pedagógica que busque ser efectiva en este entorno debe ser capaz de superar estas limitaciones, proporcionando un medio para la práctica activa y frecuente que compense la escasez de exposición formal e informal.

Se aborda la relevancia de teorías pedagógicas como el construccionismo y el aprendizaje activo, se analiza el papel de la IA en la motivación y la adaptabilidad del aprendizaje, se profundiza en los aspectos lingüísticos clave, y se justifica la integración de herramientas como GTM. Este marco proporciona la fundamentación para diseñar e implementar una propuesta didáctica basada en la programación de modelos de clasificación de audio con GTM, articulada con prácticas fonéticas para el reconocimiento y la producción de pares mínimos en inglés del nivel A1 y A2 del MCER.

5.1 Fundamentos pedagógicos: construccionismo y aprendizaje activo

5.1.1. *Construccionismo y aprendizaje activo*

El construccionismo, teoría educativa propuesta por Seymour Papert y fundamentada en el constructivismo de Jean Piaget, sostiene que el aprendizaje más significativo ocurre cuando los estudiantes participan activamente en la construcción de artefactos concretos y compartibles

(Badilla Saxe y Chacón Murillo, 2004). Papert plantea que el proceso de ensayo-error y corrección, conducen a la creación y el aprendizaje, y subraya la responsabilidad de la sociedad y la cultura en proveer los recursos y condiciones adecuadas para que los aprendices puedan construir sus entidades públicas y, en el proceso, su propio conocimiento (Badilla Saxe y Chacón Murillo, 2004).

Este enfoque se alinea directamente con el aprendizaje activo, definido como una metodología pedagógica centrada en el estudiante, donde este asume un rol protagónico en su proceso educativo, participando activamente en la construcción de su propio conocimiento a través de experiencias prácticas y la interacción con su entorno (Su y Zhong, 2022; Valente, 2023). Consiste en aprender "haciendo" mediante actividades prácticas y la reflexión, que conectan la teoría con situaciones del mundo real, fomentando la participación significativa y las discusiones constructivas, así como el trabajo colaborativo (Koneru, 2023; Su y Zhong, 2022; Valente, 2023). La integración de GTM, donde los aprendices participan en la recolección de datos, el etiquetado de audios y el entrenamiento de modelos, introduce un componente constructor que fortalece la alfabetización digital y el pensamiento computacional (Carney et al., 2020; Dwivedi et al., 2024; Sayginer y Tekdal, 2023). Esta perspectiva es fundamental para el diseño de la propuesta didáctica del estudio, que busca involucrar activamente a los estudiantes en la creación y experimentación con modelos de IA para mejorar su pronunciación.

5.1.2. Potencial de la IA en la motivación y el rendimiento

La IA se ha consolidado como un recurso clave en la enseñanza de EFL, especialmente en el entrenamiento de la pronunciación, gracias al desarrollo de tecnologías como el ASR y la TTS (Tejedor-García et al., 2020; Aryanti y Santosa, 2024; Zhao, 2024). Estas herramientas no solo favorecen la precisión fonética, sino que también ofrecen retroalimentación inmediata y adaptativa, lo que potencia la personalización del aprendizaje y la motivación de los estudiantes (Silva et al., 2024; López-Minotta et al., 2025).

Los sistemas CAPT han demostrado resultados significativos en diferentes contextos. Tejedor-García et al. (2020) reportaron que el sistema de reconocimiento de voz de Google alcanzó un nivel de precisión del 95 %, cercano al desempeño humano. En sus experimentos con estudiantes hispanohablantes de nivel intermedio, aquellos que utilizaron ASR y TTS incrementaron su rendimiento en discriminación y producción de sonidos, con mejoras relativas

del 205 % frente al 82 % obtenido con instrucción tradicional. Además, los aprendices con menor dominio inicial aprovecharon de forma más intensa la TTS, logrando avances notables en reconocimiento y producción.

Resultados similares se han documentado en otros escenarios. Neri et al. (2008) demostraron que niños de 11 años mejoraron la calidad segmental de 11 fonemas problemáticos gracias a la retroalimentación automática. López-Minotta et al. (2025), en un estudio con estudiantes de quinto grado, evidenciaron que la puntuación media de pronunciación pasó de 6.40 a 13.20 y la expresión oral de 3.28 a 7.12 tras 16 semanas de intervención, confirmando el potencial de la IA en educación primaria. En la educación superior, Senowarsito y Ardini (2023) encontraron que el uso de la aplicación ELSA Speak incrementó la media de pronunciación de 66.06 a 85.63 en nueve sesiones, con una correlación positiva entre el entrenamiento con IA y el aprendizaje autónomo.

Analizando profundamente los resultados más allá del área cuantitativa, los hallazgos demuestran que la retroalimentación automática contribuye a la motivación y a la confianza del estudiante. La inmediatez y especificidad de la retroalimentación permiten corregir hipótesis previas, promueven un procesamiento más profundo y reducen el temor a cometer errores, favoreciendo el aprendizaje autodirigido (Ashwell y Elam, 2017; Leow, 2019; Maulana y Suharto, 2024; Sabaruddin et al., 2024). Si bien pueden existir errores en la retroalimentación generada por los CAPT, los estudios coinciden en que su impacto global es positivo para el desarrollo de la pronunciación y la motivación de los aprendices (Neri, 2007).

En conjunto, estas investigaciones demuestran que la IA no solo mejora el rendimiento fonético en diversos niveles educativos, sino que también potencia la motivación, la autonomía y la confianza de los estudiantes. Sin embargo, la mayor parte de estas investigaciones se han llevado a cabo en contextos universitarios o con aplicaciones comerciales. Aún existe un vacío en la aplicación de enfoques constructivistas con herramientas accesibles como GTM, especialmente en la enseñanza de pares mínimos en educación primaria, lo que refuerza la pertinencia de la presente investigación.

5.1.3. Rol docente y andamiaje en entornos con IA

El papel del docente en un entorno de aprendizaje mejorado por IA es fundamentalmente el de un facilitador y guía, en lugar de un mero transmisor de información (Baez et al., 2023; Okagbue et al., 2023). La IA, en este contexto, debe ser vista como una herramienta que apoya y complementa las metodologías instruccionales eficaces, no como un reemplazo de los educadores humanos (Duisenova y Zhorabekova, 2024; Forero-Corba y Negre Bennasar, 2024).

El andamiaje, entendido como una forma de asistencia o mediación proporcionada al estudiante para que pueda alcanzar metas (Gerson, 2006), es crucial, y los sistemas de IA pueden ayudar a los docentes proporcionando retroalimentación personalizada y continua (Okagbue et al., 2023). Por ejemplo, los sistemas CAPT ofrecen retroalimentación correctiva y evaluación automática de la calidad de la pronunciación, complementando la instrucción en el aula (Tejedor-García, 2020). Esta retroalimentación, incluso si es solo un pequeño cambio en la pedagogía, puede resultar en un aprendizaje estudiantil mejorado (Bejtíc, 2023). La gestión de errores es un aspecto vital, ya que los docentes pueden utilizar las herramientas de IA para identificar las dificultades reales de los estudiantes, como los sonidos o actividades más difíciles en el aprendizaje de un idioma (Tejedor-García, 2020). Además, los profesores capaces de usar IA para la enseñanza pueden mejorar su efectividad e influir positivamente en la motivación y autoeficacia de los estudiantes (Baez et al., 2023). Para una implementación exitosa, los planes de estudio deben incorporar transversalmente los conceptos de las nuevas tecnologías inteligentes, y los educadores necesitan capacitación y desarrollo profesional para integrar eficazmente la IA en sus prácticas (Forero-Corba y Negre Bennasar, 2024; Rizvi et al., 2023; Silva et al., 2024). Los sistemas de IA también pueden ofrecer apoyo emocional e identificar errores en los diálogos, complementando el papel del tutor humano (Ortega-Ochoa, 2022).

5.2. Enfoque lingüístico y categorías de la investigación

5.2.1. Categorías centrales de la investigación

La investigación se centra en la competencia fonética, abarcando dos variables de interés: el reconocimiento auditivo y la producción oral. A continuación, se definen y sustentan las competencias evaluadas en el contexto de la investigación:

5.2.1.1 Reconocimiento auditivo. Se refiere a la capacidad de los estudiantes para discriminar y procesar los contrastes fonológicos en los pares mínimos, incluyendo tanto la percepción de los sonidos como la construcción de significado a partir de ellos (Pennington y Rogerson-Revell, 2019). La mejora del reconocimiento auditivo se correlaciona con una mejor producción oral (Acosta Acosta, 2023), lo que subraya la importancia de los ejercicios fonéticos dirigidos. Investigaciones sugieren que la percepción del habla de una L2, moldeada por la experiencia del primer idioma, o materna (L1), se desarrolla a través de la capacidad de los aprendices para discernir las diferencias acústicas entre los sonidos de su L1 y L2 (Mora et al., 2014; Iverson y Evans, 2009).

5.2.1.2 Producción oral. Es la habilidad de pronunciar con inteligibilidad y precisión los pares mínimos, integrando tanto la corrección fonética como la claridad comunicativa (Acosta Acosta, 2023). El éxito de la comunicación en un idioma extranjero depende en gran medida de la inteligibilidad, la comprensión, el acento y la fluidez del habla (Tejedor-García, 2020). La inteligibilidad, definida como la capacidad de ser entendido, es un objetivo pedagógico más adecuado que la reducción del acento (Pennington y Rogerson-Revell, 2019). La capacitación en producción a menudo genera mejoras más fuertes o completas en comparación con la capacitación centrada solo en la percepción (Acosta Acosta, 2023).

5.2.2. Teoría fonética, pares mínimos y conciencia fonológica

El español y el inglés presentan diferencias significativas en sus patrones fonológicos, incluyendo la correspondencia entre sonido y letra, el tamaño del inventario de fonemas, la

realización alofónica de sonidos y el lugar y modo de articulación (Silva Valencia, 2022; Uribe-Enciso et al., 2019). Estas diferencias soportan el uso de estrategias pedagógicas específicas y diferenciadas entre idiomas.

Los pares mínimos son dos palabras que se distinguen en significado por el cambio de un solo segmento sonoro en la misma posición (Acosta Acosta, 2023; Asmaul Husna et al., 2023). La práctica con pares mínimos es una técnica valiosa en la enseñanza y aprendizaje de la pronunciación, ya que están presentes en casi todos los idiomas y pueden contrastar los sonidos de la L1 y la L2. Un sistema CAPT que incorpora pares mínimos tiene como objetivo mejorar la pronunciación en L2 a nivel segmental (Tejedor-García, 2020). Estudios indican que el entrenamiento perceptivo basado en pares mínimos puede ser eficaz y facilitar el desarrollo de habilidades productivas (Carlet, 2017).

Entre los contrastes críticos para hispanohablantes aprendiendo inglés se encuentran:

5.2.2.1 Contrastes vocálicos críticos. Los pares /i:~ɪ/ y /u:~ʊ/ son particularmente desafiantes debido a que el sistema vocálico del español es más reducido, con cinco vocales, en comparación con el inglés, con quince vocales (Silva Valencia, 2022; Valenzuela y French, 2023). Esta disparidad fonológica lleva a que los hispanohablantes no perciban o no produzcan con precisión las distinciones vocálicas del inglés (Silva Valencia, 2022).

5.2.2.2 Contrastes consonánticos críticos. Los hispanohablantes comúnmente asimilan las fricativas sonoras del inglés a sus contrapartes sordas en español y enfrentan dificultades con fricativas interdentalas como /ð/ y /θ/ (Acosta Acosta, 2023; Uribe-Enciso et al., 2019). La retroalimentación proporcionada por la IA puede reforzar la instrucción teórica y la producción correcta de estas fricativas (Neri, 2007).

La aplicación de protocolos para la elaboración de listas de pares mínimos, que considera la L1 y L2 del participante y las tecnologías ASR y TTS, ha demostrado ser útil y efectiva en el contexto de las herramientas CAPT (Tejedor-García, 2020).

5.2.3. El “Noticing” en la adquisición de L2

La Hipótesis del Noticing de Schmidt subraya la importancia fundamental de la atención y la conciencia en el proceso de aprendizaje de una L2 (Leow, 2019). Esta hipótesis es considerada uno de los pilares teóricos más influyentes en la adquisición de L2, centralizando el papel de la atención y la conciencia en el proceso de aprendizaje. El noticing se define como el registro de la ocurrencia de un evento estímulo en la conciencia y su posterior almacenamiento en la memoria a largo plazo (Leow, 2019). Para adquirir cualquier característica lingüística de la L2 es imprescindible que dicha característica sea notada en el aprendizaje de la L2, lo que requiere una atención focalizada y, al menos, un bajo nivel de conciencia. Se ha demostrado que una mayor conciencia y una mayor profundidad de procesamiento están directamente relacionadas con la solidez del aprendizaje (Leow, 2019).

La fuerte influencia de la L1 puede ser tan abrumadora que los aprendices a menudo no logran notar las discrepancias entre su propia producción y los sonidos objetivo de la L2 (Neri, 2007; Neri et al., 2002; Neri et al., 2008). De hecho, los errores de pronunciación en L2 rara vez son intentos aleatorios, sino que suelen reflejar el sistema fonológico de la L1 del aprendiz, a menudo combinado con la transferencia ortográfica (Pennington y Rogerson-Revell, 2019).

En este contexto, la retroalimentación externa es fundamental para el aprendizaje de la pronunciación, ya que ayuda a los estudiantes a tomar conciencia de estas discrepancias y a superarlas (Tejedor-García, 2020). Es crucial que esta retroalimentación no dependa únicamente de las percepciones del propio estudiante para que sea efectiva en la creación de esta conciencia, dado que la influencia de la L1 puede cegar al aprendiz ante sus propios errores (Neri et al., 2002; Neri, 2007). La retroalimentación permite a los estudiantes enfocar su atención en problemas individuales específicos y los estimula a buscar la auto-mejora; un tipo de retroalimentación que no se basa en las propias percepciones del estudiante es necesario para abordar los errores donde la influencia de la L1 es abrumadora (Neri, 2007).

Los sistemas ASR pueden proporcionar retroalimentación correctiva inmediata, lo que es beneficioso para mejorar la precisión de la pronunciación (Ashwell y Elam, 2017). Cuando un sistema ASR no logra transcribir con precisión la producción oral de un aprendiz no nativo, esto puede señalar un error de pronunciación potencial, lo que lleva al aprendiz a analizar y reformular su enunciado de una manera más comprensible (Ashwell y Elam, 2017; Lai y Chen, 2022). Es

crucial que la retroalimentación no dependa únicamente de las percepciones del propio estudiante para que sea efectiva en la creación de esta conciencia (Tejedor-García, 2020). Las herramientas basadas en IA ofrecen experiencias de aprendizaje personalizadas, adaptativas y atractivas, apoyando el aprendizaje autónomo (López-Minotta et al., 2025; Sabaruddin et al., 2024; Tejedor-García, 2020). La retroalimentación personalizada y exagerada (aumentando la duración del sonido o el movimiento de la animación) puede reforzar significativamente la capacidad de percepción y, consecuentemente, la habilidad de producción del habla, haciendo que las distinciones sean más memorables (Bu et al., 2021). Además, la retroalimentación correctiva debe ser clara, explícita y permitir la autorreparación de los errores fonéticos (Neri, 2007).

5.2.4. Percepción del estudiante en procesos mediados por IA

La percepción del estudiante es una categoría fundamental en el contexto de los procesos de aprendizaje mediados por la IA. Los participantes en estudios previos han valorado positivamente los sistemas TTS, indicando que les ayudaron a darse cuenta de las diferencias en los pares mínimos y a practicar palabras difíciles, lo que denota una alta utilidad percibida (Tejedor-García, 2020).

La IA en educación mejora la personalización de la enseñanza, ofreciendo retroalimentación en tiempo real y adaptando el contenido y las estrategias pedagógicas a las necesidades individuales (Silva et al., 2024). La capacidad de los alumnos para entrenar "en cualquier momento y en cualquier lugar, a su propio ritmo" (Tejedor-García, 2020) fomenta la autonomía. El éxito en la mejora de la pronunciación y la reducción del miedo a cometer errores, así como el aumento de la confianza (Maulana y Suharto, 2024; Sabaruddin et al., 2024), contribuyen a la competencia percibida. A pesar de desafíos como el limitado conocimiento tecnológico y la falta de acceso a recursos, la implementación de la IA tiene el potencial de ofrecer experiencias de aprendizaje personalizadas, adaptativas y atractivas (Silva et al., 2024). La propuesta didáctica incluye encuestas al final de cada sesión para recopilar datos cualitativos sobre la motivación, la competencia percibida, el deseo de continuar practicando, la ayuda del modelo para la escucha y el trabajo en equipo, además de comentarios sobre dificultades y soluciones. Estos datos son cruciales para medir la mejora de la competencia fonética, considerando tanto los resultados cuantitativos como la experiencia subjetiva de los estudiantes.

5.3 Integración tecnológica y fundamentación de la propuesta

5.3.1 Machine Learning y Teachable Machine

El ML es una rama de la IA que se enfoca en enseñar a los sistemas a aprender de los datos para resolver problemas a partir de ejemplos o experiencias previas (Forero-Corba y Negre Bannasar, 2024; Okagbue et al., 2023). Este campo está impactando todas las áreas del conocimiento, incluyendo la educación (Forero-Corba y Negre Bannasar, 2024).

GTM se presenta como una herramienta pertinente y accesible que permite a estudiantes de primaria interactuar con el ML de forma práctica (Sayginer y Tekdal, 2023). GTM promueve el aprendizaje continuo y ha demostrado ser adecuada para enseñar principios de ML a niños pequeños, implementándose en diversas configuraciones de K-12 (Forero-Corba y Negre Bannasar, 2024; Sayginer y Tekdal, 2023; Pujari et al., 2022). El aspecto más relevante de GTM para esta investigación es su componente constructorista, donde los aprendices participan directamente en la recolección de datos de audio, su etiquetado y el entrenamiento de los modelos de clasificación (Carney et al., 2020; Sayginer y Tekdal, 2023). Esta metodología no solo facilita la práctica fonética, sino que involucra al estudiante en los principios del ML, fortaleciendo la alfabetización digital y el pensamiento computacional (Dwivedi et al., 2024). Sin embargo, Sayginer y Tekdal (2023) subrayan que, a pesar del potencial educativo de GTM, su uso sigue siendo limitado y la investigación sobre la herramienta se encuentra en etapas iniciales, concentrándose principalmente en la enseñanza de conceptos de ML en contextos K-12. Este panorama revela un vacío claro en la literatura en cuanto al empleo de GTM para la práctica fonética, justificando la intervención didáctica propuesta por el estudio.

5.3.2. Comparación de enfoques: tradicional vs. mediado por IA

La integración de la IA en la enseñanza de la pronunciación ofrece múltiples ventajas frente a los métodos tradicionales:

5.3.2.1. Retroalimentación inmediata y personalización. La IA puede ofrecer retroalimentación instantánea sobre la pronunciación, gramáticas, vocabulario y estructura de la oración. Esta inmediatez permite a los estudiantes identificar y corregir errores en tiempo real, facilitando un progreso más rápido y una comprensión más profunda (Neri, 2007; Zhao, 2024).

5.3.2.2. Entorno de baja ansiedad. La tecnología puede crear un ambiente de aprendizaje libre de estrés y privado, donde los estudiantes se sienten más cómodos practicando sin el miedo a cometer errores frente a sus compañeros o profesores (Neri, 2007; Senowarsito y Ardini, 2023; Tejedor-García, 2020).

5.3.2.3. Accesibilidad e inclusión. Las plataformas de aprendizaje de idiomas impulsadas por IA pueden adaptarse a diversos estilos de aprendizaje, preferencias y habilidades, con características como texto a voz y contenido multimodal, lo que las hace más accesibles (Guralnick et al., 2023).

No obstante, existen limitaciones en el enfoque mediado por IA:

5.3.2.4. Precisión del ASR. Los sistemas ASR pueden no transcribir con precisión el habla no nativa o rechazar pronunciaciones inteligibles, lo que puede confundir al aprendiz (Neri, 2007; Neri et al., 2002). La efectividad del ASR puede degradarse en entornos ruidosos y requiere grandes volúmenes de datos para mejorar la precisión (Forero-Corba y Negre Bennasar, 2024; Pujari et al., 2022).

5.3.2.5. Calidad de la retroalimentación. Aunque útil, la retroalimentación a veces puede ser insuficiente, difícil de entender o no indicar cómo corregir el error (Forero-Corba y Negre Bennasar, 2024).

Es importante destacar que las aplicaciones de IA en la enseñanza de idiomas suelen complementar los enfoques tradicionales en lugar de reemplazarlos por completo, siendo la combinación de la interacción humana con las capacidades tecnológicas la más efectiva (Aryanti y Santosa, 2024; Neri, 2007; Tejedor-García, 2020; Zhao, 2024). Esta investigación se posiciona

en esta complementariedad, utilizando la IA como una herramienta que enriquece y apoya la instrucción tradicional.

5.3.3. Competencias digitales y pensamiento computacional

La programación de modelos de IA, incluso a través de interfaces amigables como GTM, fomenta el desarrollo de competencias digitales esenciales:

5.3.3.1. Alfabetización digital. La interacción práctica con herramientas como GTM ayuda a los estudiantes a comprender tecnologías innovadoras como la IA y el ML a una edad temprana, contribuyendo al desarrollo de habilidades tecnológicas y aumentando su alfabetización digital (Sayginer y Tekdal, 2023; Su et al., 2023).

5.3.3.2. Recolección y análisis de datos. Al crear sus propios proyectos con GTM, los estudiantes aplican habilidades como la recopilación y el análisis de datos, la clasificación y la interpretación (Sayginer y Tekdal, 2023). Estas competencias son claves para la transversalización curricular de TIC e IA en primaria.

Esta democratización del acceso a la IA permite que estudiantes con dispositivos menos potentes puedan probar y experimentar con sus aplicaciones (Carney et al., 2020; Sayginer y Tekdal, 2023). La integración de la IA en la educación es una innovación crucial que tiene el potencial de cerrar la brecha digital y escolar, preparando a los estudiantes para los desafíos de la sociedad posGTModerna y el mundo digital actual (Forero-Corba y Negre Bennasar, 2024; Su et al., 2023).

5.3.4. Limitaciones y consideraciones éticas del uso de IA

El uso de la IA en la educación, aunque prometedor, conlleva importantes riesgos y retos que deben ser abordados con un enfoque equilibrado:

5.3.4.1. Consideraciones éticas y privacidad de datos. Los sistemas de IA a menudo requieren grandes volúmenes de datos personales para funcionar eficazmente, lo que plantea importantes preocupaciones éticas y de privacidad que deben gestionarse con precaución (Silva et al., 2024; Zhao, 2024). El Big Data, fundamental para las técnicas de ML, debe mantener la integridad ética y moral de la información de los participantes y el currículo (Forero-Corba y Negre Bennasar, 2024).

5.3.4.2. Sesgos en los datos. Si los datos de entrenamiento de los modelos de IA no están bien representados, pueden introducirse sesgos que afecten negativamente la precisión y equidad de los sistemas (Zhao, 2024).

5.3.4.3. Comprensión limitada del contexto y la inteligencia emocional. La IA todavía tiene una comprensión limitada del uso del lenguaje, el sarcasmo, el contexto cultural y la inteligencia emocional en comparación con un profesor humano, lo que puede resultar en una menor interacción humana (Zhao, 2024).

5.3.4.4. Brecha digital. Es fundamental fomentar una alfabetización digital crítica para que los estudiantes comprendan los riesgos y sesgos de la IA, lo que implica no solo usar la tecnología, sino también reflexionar críticamente sobre sus implicaciones personales y sociales (Forero-Corba y Negre Bennasar, 2024; Silva et al., 2024).

5.3.4.5. Limitaciones técnicas de ASR. Como se mencionó, el ASR puede presentar problemas de precisión con el habla no nativa o en entornos ruidosos, y la retroalimentación puede no ser siempre lo suficientemente explícita o fácil de entender (Neri, 2007).

Para una implementación exitosa, se requiere un enfoque cuidadoso y equilibrado que combine la innovación tecnológica con el valor irremplazable del toque humano en la educación (Forero-Corba y Negre Bennasar, 2024; Okagbue et al., 2023; Silva et al., 2024; Zhao, 2024). La IA debe ser un factor que impulse la calidad metodológica e instrumental de las metodologías instruccionales tradicionales, no ser un factor destructivo que busque su reemplazo.

5.4. Síntesis y pertinencia para la investigación

La presente investigación aborda el efecto de una propuesta didáctica sobre el reconocimiento auditivo y la producción oral de pares mínimos en inglés en estudiantes de primaria. Se integra el construccionismo y el aprendizaje activo mediante la programación práctica de modelos de clasificación de audio en GTM, donde los estudiantes graban, entrenan y prueban modelos, recibiendo retroalimentación inmediata sobre la precisión de su pronunciación. Este proceso iterativo, que valora los errores como oportunidades de aprendizaje, refuerza el enfoque construccionista y alinea directamente con el objetivo de diseñar e implementar una propuesta didáctica que fomente estas habilidades.

La eficacia de la IA en la motivación y el aprendizaje, evidenciada por la capacidad de ofrecer retroalimentación personalizada y crear entornos de baja ansiedad, es central para el éxito de la intervención. Las diferencias fonológicas entre el español y el inglés, particularmente en los pares mínimos críticos, subrayan la necesidad de una instrucción focalizada. La Hipótesis del Noticing de Schmidt y la importancia de la retroalimentación externa, a menudo mejorada por la IA, justifican el diseño de actividades que promuevan la conciencia fonológica y la discriminación auditiva. La percepción positiva del estudiante hacia las herramientas de IA, junto con el fomento de la autonomía y la confianza, valida la aproximación tecnológica de la propuesta.

Al mismo tiempo, la investigación reconoce las limitaciones de las tecnologías de herramientas como GTM, como la precisión del ASR en ambientes ruidosos y la necesidad de grandes volúmenes de datos, así como las consideraciones éticas sobre la privacidad y los sesgos. Esto resalta la importancia de una implementación reflexiva y complementaria de la IA, sin desplazar el rol insustituible del docente. El estudio busca medir la mejora en el reconocimiento y la producción de pares mínimos, empleando pruebas y rúbricas validadas para el GE y de control. La integración de los antecedentes sobre la consolidación de la IA en la enseñanza de la pronunciación, la evidencia de resultados significativos con CAPT, y el vacío identificado en la literatura sobre el uso de GTM para la práctica fonética, proporciona un sólido sustento empírico y conceptual para la metodología y los objetivos de esta investigación, asegurando que la propuesta no solo sea innovadora sino también pertinentemente fundamentada en el contexto educativo actual.

6 Metodología

6.1 Enfoque de investigación

Esta investigación adopta un diseño mixto de triangulación concurrente. El componente cuantitativo se estructura mediante un diseño cuasiexperimental con pretest – postest y grupo control (GC) no equivalente, orientado a evaluar el efecto causal de la intervención constructorista mediada por GTM sobre la producción oral de pares mínimos en inglés. El componente cualitativo se apoya en una metodología fenomenológica descriptiva, diseñada para indagar en las percepciones, motivaciones y experiencias de aprendizaje de los estudiantes durante la intervención.

6.2 Diseño de investigación

El diseño adoptado es cuasiexperimental de tipo pretest y postest con grupo de control no equivalente. Este diseño se ajusta al contexto escolar, donde los grupos de quinto grado ya están conformados y no es posible la asignación aleatoria de participantes. La aplicación de los mismos instrumentos en condiciones controladas garantiza la comparabilidad entre los grupos y la validez de los resultados.

El GE recibe instrucción fonética complementada con la estrategia mediada por GTM, mientras que el grupo de control continúa con el modelo tradicional basado en explicación y repetición. Ambos grupos acceden a la misma enseñanza teórica, pero la intervención del GE incorpora programación de modelos de clasificación de audio en GTM, lo que permite a los estudiantes grabar, entrenar y validar sus producciones.

6.3 Contexto y participantes

La aplicación del presente estudio se realiza en el INEM José Celestino Mutis, una institución educativa pública ubicada en Armenia, Quindío (Colombia). Esta institución atiende población de diferentes estratos socioeconómicos, representativa del contexto educativo colombiano donde se implementan políticas de bilingüismo con recursos limitados. La muestra se

compone de 44 estudiantes de quinto grado de primaria, seleccionados mediante muestreo intencional por conveniencia basado en la disponibilidad de grupos intactos y la factibilidad de implementación de la intervención. Las características demográficas de los participantes son las siguientes:

- Rango de edad: 9-12 años ($M = 10.4$, $DE = 0.9$)
- Distribución por género: 52.3 % femenino ($n = 23$), 47.7 % masculino ($n = 21$)
- Nivel socioeconómico: Estratos 1 a 4 (según clasificación del DANE)
- Lengua materna: Español (100 % de los participantes)
- Experiencia previa en inglés: Limitada a instrucción escolar básica (1-2 horas semanales)

La asignación a los grupos se realiza mediante grupos intactos preexistentes para minimizar la disrupción del proceso educativo institucional:

- GE: 23 estudiantes iniciales (22 completaron la intervención)
- GC: 21 estudiantes

6.3.1 Criterios de inclusión

1. Estar matriculado en quinto grado de la institución participante.
2. Ser hablante nativo de español.
3. Contar con consentimiento informado de padres/acudientes y asentimiento del estudiante.

6.3.2 Criterios de exclusión

1. Ausencias superiores al 50 % de las sesiones de intervención.
2. Presentar alguna discapacidad que le impida el uso de GTM.
3. No completar las evaluaciones pretest o postest.

6.4 Variables y categorías de estudio

La investigación considera una variable independiente y tres dependientes:

6.4.1 *Variable independiente.*

6.4.1.1 Propuesta didáctica con GTM. Corresponde a la intervención pedagógica centrada en el aprendizaje activo y el construccionismo, mediante la programación de modelos de clasificación de audio en GTM. Esta variable integra las TIC e IA en el currículo de primaria como una estrategia innovadora para la enseñanza de la pronunciación.

6.4.2 *Variables dependientes*

6.4.2.1 Reconocimiento auditivo de pares mínimos en inglés. Se refiere a la capacidad de los estudiantes para discriminar auditivamente contrastes fonológicos críticos para hispanohablantes. Esta variable está asociada con las dimensiones de discriminación y conciencia fonológica.

6.4.2.2. Producción oral de pares mínimos en inglés. se refiere al nivel de precisión articuladora e inteligibilidad alcanzado por los estudiantes en la pronunciación de pares mínimos. Se articula con la producción oral como componente central de la expresión lingüística.

Cada variable se operacionaliza mediante indicadores cuantitativos: medias (M), desviaciones estándar (DE), diferencias pre-post y entre grupos analizadas con pruebas t , significancia estadística (p), tamaños del efecto (d, g) e intervalos de confianza al 95 %. Asimismo, tanto en el reconocimiento auditivo como en la producción oral se realiza un análisis desagregado por categorías fonológicas con el fin de precisar el desempeño de los estudiantes en contrastes específicos. Esta categorización permite identificar patrones de dificultad diferenciados y examinar la efectividad de la intervención según la naturaleza articuladora de los pares mínimos trabajados.

6.4.2.3 Percepción frente a la estrategia. Entendida como el nivel de interés, utilidad percibida y facilidad de uso que los estudiantes manifestaron frente a la propuesta didáctica. Esta variable se mide a través de encuestas tipo Likert y observaciones sistemáticas. Permite evaluar la motivación y la aceptación de la estrategia por parte de los estudiantes. De esta manera, el estudio integra dimensiones cognitivas, como el reconocimiento auditivo y la producción fonética, y afectivo-motivacionales, como la percepción, lo que posibilita un análisis integral del impacto de la propuesta didáctica.

6.5 Instrumentos de recolección de datos

6.5.1 Prueba de reconocimiento auditivo y producción oral de pares mínimos.

Este instrumento consta de 30 ítems distribuidos en dos secciones. La Sección A incluye 20 ítems de reconocimiento auditivo con grabaciones del *Oxford Learner's Dictionary*. La Sección B contiene 10 ítems de producción oral evaluados con rúbrica de tres niveles, , con una puntuación máxima posible de 20 puntos:

- Nivel 3 (Competente): Producción clara y precisa del contraste fonológico, completamente inteligible. Equivale a 2 puntos.
- Nivel 2 (En desarrollo): Producción parcialmente correcta con ligeras imprecisiones que no comprometen significativamente la inteligibilidad. Equivale a 1 punto.
- Nivel 1 (Inicial): Producción imprecisa o incorrecta del contraste que compromete la inteligibilidad. Equivale a 0 puntos.

El mismo instrumento se aplica en pretest y posttest, asegurando comparabilidad. Las producciones orales se graban digitalmente en formato m4a (48 kHz, 128 kbps) en ambiente controlado utilizando micrófono direccional para minimizar ruido ambiental. Cada estudiante lee los ítems individualmente con un intervalo de 3 segundos entre producciones.

6.5.2 Encuesta de percepción.

Se trata de una encuesta tipo Likert de cinco puntos, aplicada al GE para evaluar aspectos como la diversión, la competencia percibida, el interés en continuar practicando, la utilidad del modelo y el trabajo en equipo.

6.5.3 Observación docente.

Se utiliza una lista de cotejo durante las sesiones con indicadores de participación, persistencia, colaboración, enfoque, frustración e iniciativa. Este instrumento brinda evidencia cualitativa complementaria sobre la motivación y la dinámica de trabajo en el aula.

6.5.4 Registros de desempeño en GTM

Se recopilan las tablas de precisión por iteración y los enlaces públicos de los modelos creados por los estudiantes. Estos registros constituyen evidencia objetiva del aprendizaje en el proceso de entrenamiento y validación de modelos de audio.

6.5.5 Validez y confiabilidad

La validez de contenido se establece mediante el juicio de ocho expertos en EFL y fonética, con coeficiente V de Aiken: Sección A (claridad = 0.937; coherencia = 0.906; pertinencia = 0.875) y Sección B (claridad = 0.875; coherencia = 0.906; pertinencia = 0.875). Todos los valores superan 0.80, confirmando la validez de los ítems (Merino-Soto, 2023).

La confiabilidad interna de la Sección A se evalúa con KR-20 en una prueba piloto con seis estudiantes de sexto grado, alcanzando 0.743, lo que indica consistencia aceptable (Báez et al., 2023; Chandran et al., 2024).

La confiabilidad intercalificador (ICC) de la Sección B se calcula con el coeficiente Kappa de Cohen, obteniendo un valor de 0.66 en la prueba piloto, que representa un acuerdo sustancial entre evaluadores (Mao y Wang, 2012). Este procedimiento se mantiene en pretest y postest.

6.6 Procedimiento

El desarrollo metodológico se organiza en cuatro fases secuenciales, garantizando la coherencia entre la preparación, la intervención y la evaluación final.

6.6.1. Fase 1: *Pretest* (28 de julio – 4 de agosto)

Se aplica el instrumento inicial para establecer la línea base en ambos grupos.

6.6.1.1 Reconocimiento auditivo. Los estudiantes escuchan grabaciones de pares mínimos tomadas del Oxford Learner's Dictionary, con el fin de garantizar consistencia y neutralidad en la pronunciación.

6.6.1.2 Producción oral. Los estudiantes pronuncian pares mínimos, cuyas respuestas son grabadas y evaluadas por dos calificadores independientes, docente titular y evaluador externo.

Las condiciones de aplicación son controladas: mismo salón, computador, pantalla y sistema de sonido.

6.6.2. Fase 2: *Intervención didáctica* (6, 11, 13 y 20 de agosto)

Se desarrollan cuatro sesiones de 80 minutos con el GE, mientras el grupo de control recibe instrucción fonética tradicional. Cada sesión incluye tres momentos:

1. 15 minutos de teoría fonética: explicación de contrastes fonológicos, apoyada en presentaciones de clase.
2. 45 minutos de práctica en GTM: grabación de muestras, entrenamiento de modelos multiclase y validación de exactitud.
3. 15 minutos de pruebas y retroalimentación cruzada: verificación de los modelos con palabras nuevas y pruebas entre equipos.

Los contrastes trabajados son:

- Sesión 1: fundamentos de fonética + modelo demo en GTM.
- Sesión 2: vocales largas–cortas (/i:–ɪ/, /u:–ʊ/).
- Sesión 3: oclusivas sordas–sonoras (/p–b/, /t–d/, /k–g/).
- Sesión 4: fricativas (/f–v/, /s–z/, /θ–ð/).

Evidencias recolectadas en cada sesión:

- Enlaces de modelos (URL pública).
- Tablas de iteración vs. exactitud.
- Encuestas de percepción (escala Likert 1–5) y una pregunta abierta para comentarios.
- Observación docente (participación, persistencia, colaboración, enfoque, frustración, iniciativa).

6.6.3. Fase 3: Posttest (25 – 29 de agosto)

Se aplica nuevamente el mismo instrumento en ambos grupos para medir el progreso en el reconocimiento auditivo y la producción oral. Adicionalmente, se aplica la encuesta de percepción.

6.6.4. Fase 4: Análisis y sistematización

En esta fase se consolidan y analizan los datos recolectados a través de los diferentes instrumentos. El análisis se organiza en dos enfoques complementarios: cuantitativo y cualitativo. En el enfoque cuantitativo, los resultados de las pruebas de reconocimiento auditivo y producción oral se procesan mediante estadística descriptiva (medias, desviaciones estándar y porcentajes) e inferencial. Se aplican pruebas *t* de Welch para muestras relacionadas, a fin de comparar los puntajes del pretest y posttest dentro de cada grupo, y pruebas *t* para muestras independientes, con el propósito de contrastar los resultados finales entre el GE y el de control. Se adopta un nivel de significancia de $p < 0.05$ (Chandran et al., 2024; Shehata, 2024; Acosta Acosta, 2023). La magnitud de los cambios se interpreta mediante tamaños del efecto (Cohen's *d* para diseños pareados y Hedges' *g* para comparaciones entre grupos), junto con intervalos de confianza del 95 % para

estimar la precisión de los resultados. Estos indicadores complementan el valor de significancia y permiten una interpretación más robusta de los hallazgos.

En el enfoque cualitativo, se categorizan las respuestas abiertas de las encuestas de percepción y las observaciones docentes en cuatro dimensiones:

1. Motivación y disfrute (2 ítem): Evalúa el nivel de compromiso y satisfacción durante las actividades.
2. Autoeficacia y confianza (2 ítems): Mide la percepción de competencia y seguridad en la pronunciación.
3. Usabilidad y utilidad percibida (1 ítem): Explora la facilidad de uso y valor pedagógico atribuido a GTM.
4. Experiencia constructorista: Indaga con una pregunta abierta sobre la experiencia de crear y entrenar modelos propios.

La integración de ambos enfoques posibilita una interpretación más completa del impacto de la propuesta didáctica, articulando los resultados estadísticos con la percepción y la experiencia de los estudiantes. Los resultados cuantitativos y cualitativos se triangulan para enriquecer la interpretación, integrando la evolución en reconocimiento y producción con la percepción de los estudiantes y la observación docente. Este procedimiento asegura la coherencia entre los datos recogidos y las categorías definidas en la investigación.

6.7 Consideraciones éticas

La investigación se desarrolla en estricta conformidad con los principios éticos aplicables a estudios educativos con población infantil, observando las disposiciones de la Resolución 8430 de 1993 del Ministerio de Salud de Colombia y las directrices internacionales para la protección de menores en contextos de investigación académica. Este estudio se clasifica como de *riesgo mínimo*, dado que las actividades educativas propuestas no implican afectaciones mayores a la rutina escolar habitual.

En primer lugar, se cuenta con la autorización institucional otorgada por la rectoría de la Institución Educativa INEM José Celestino Mutis, lo que asegura que la intervención se integre al plan académico sin alterar el desarrollo normal de las clases.

En segundo lugar, se garantiza el derecho a la autonomía y consentimiento informado. Para ello, los padres y acudientes firman un consentimiento en el que se explican los objetivos del estudio, la naturaleza voluntaria de la participación, la ausencia de riesgos físicos o psicológicos, y los beneficios educativos esperados. Paralelamente, los estudiantes manifiestan su asentimiento informado, expresando su acuerdo en participar de manera libre y voluntaria.

En tercer lugar, se preserva la confidencialidad y anonimato de los datos. Los registros se codifican mediante identificadores alfanuméricos, evitando cualquier relación directa con los nombres de los participantes. Las grabaciones de voz se emplean únicamente para los análisis de confiabilidad y validez, sin ser divulgadas ni utilizadas con fines ajenos a la investigación.

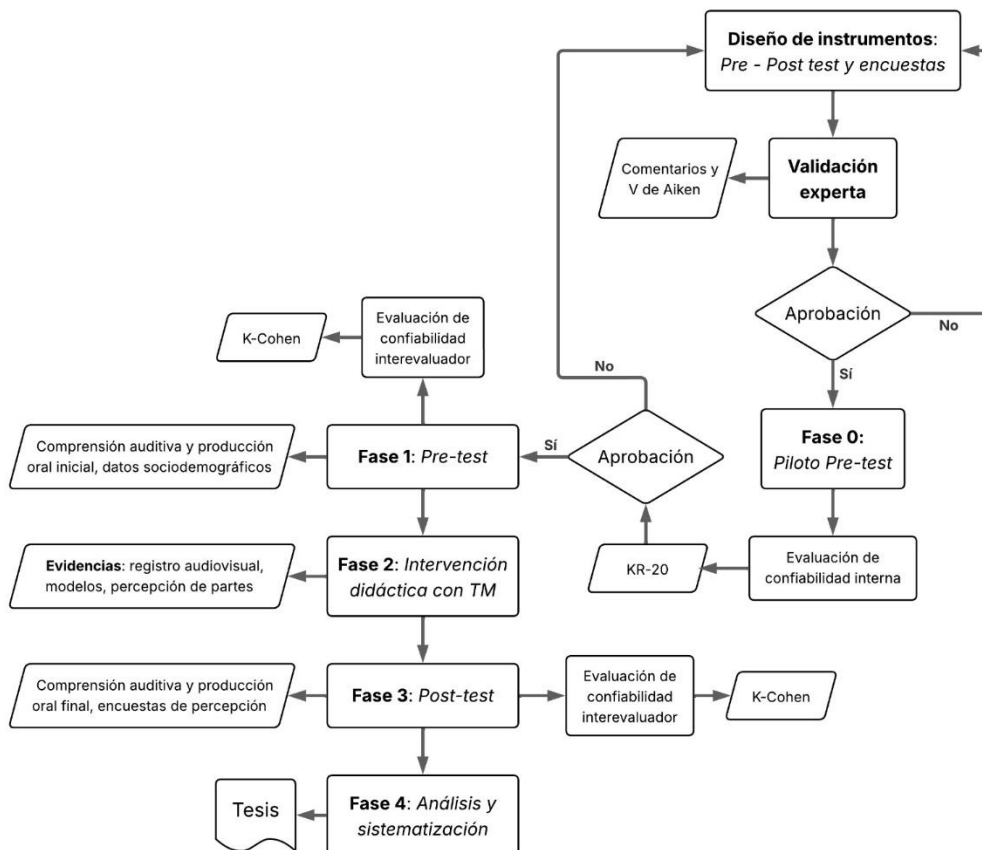
En cuarto lugar, durante las sesiones que se llevaron a cabo en las cuales los participantes entrenaron sus propios modelos de TM con diferentes parejas de pares mínimos en inglés previamente asignados a los grupos de trabajo la herramienta de IA TM se usó en línea en los equipos de la institución educativa. Como los estudiantes trabajaban en su mayoría en simultáneo y a la par las grabaciones de control de ruido se realizaron de manera grupal y los participantes guardaron silencio durante estos periodos para que el modelo fuera entrenado con el ruido externo al aula de clase. Además, como se encuentra en la página web de la herramienta, si bien los modelos son publicados en los servidores de Google, no son publicadas las muestras usadas para preparar y crear los modelos. Al ser modelos publicados online, solamente pueden acceder a ellos aquellas personas que tengan la URL, y aún en el caso de tenerla los modelos no se pueden modificar.

Finalmente, se asegura que la participación sea completamente voluntaria, gratuita y en un entorno seguro, respetando la dignidad de los estudiantes y garantizando que en ningún caso se vean afectados sus procesos académicos o personales. Con estas medidas, el estudio cumple con los estándares éticos nacionales e internacionales, asegurando el respeto a los derechos y la protección de los participantes en todas las fases del proyecto.

6.8 Integridad académica y uso de IA

En concordancia con el Acuerdo UManizales 002 (2023), se emplearon herramientas de inteligencia artificial (ChatGPT, NotebookLM, SciSpace, Gemini, Copilot y Turnitin) exclusivamente como apoyo instrumental en tareas de redacción preliminar, organización de ideas, búsqueda bibliográfica, verificación estilística y análisis de datos. Estas aplicaciones no sustituyeron la argumentación académica ni la interpretación crítica de los investigadores, quienes ejercieron control permanente sobre el proceso. Toda salida de IA fue revisada manualmente y contrastada con fuentes primarias. La trazabilidad de este uso responsable se documenta en la Bitácora de IA (Anexo 11), donde se consignan fecha, herramienta, prompt, salida generada, uso efectivo y forma de verificación.

Figura 1 Diagrama Metodológico de la Investigación



La Figura 1 sintetiza el proceso metodológico de la investigación, desde el diseño y validación de los instrumentos hasta las fases de aplicación, intervención, postest y análisis. En el diagrama se evidencian los mecanismos de validez y confiabilidad, así como las fases secuenciales del estudio, lo que asegura la coherencia entre la preparación, la intervención didáctica y la evaluación de los resultados.

7. Resultados y análisis

Este capítulo presenta los hallazgos de la intervención didáctica implementada, bajo un diseño cuasiexperimental con mediciones pretest y postest. El estudio contó con un GE ($n=22$) que participó en la intervención con GTM y un GC ($n=21$) que continuó con instrucción tradicional. Se describen primero los resultados descriptivos y la verificación de la comparabilidad inicial de los grupos. A continuación, se exponen los hallazgos para las pruebas de Sección A (reconocimiento auditivo) y Sección B (producción oral), incluyendo la verificación de supuestos estadísticos y los análisis inferenciales correspondientes. Estos análisis se potencian con el cálculo de tamaños del efecto y sus intervalos de confianza al 95 %, ofreciendo una interpretación pedagógica robusta de su magnitud. Finalmente, se integran los resultados cualitativos para ofrecer una visión completa del impacto de la intervención.

7.1 Comparabilidad inicial de los grupos

Para establecer la comparabilidad inicial entre los grupos, se realizó un análisis de línea base que confirmó la ausencia de diferencias estadísticamente significativas antes de la intervención. En la Tabla 1 se observa que tanto GE ($M = 11.77$) como GC ($M = 12.05$) partían de un desempeño muy cercano en la Sección A, diferencia que la prueba t de Welch corroboró como no significativa. De manera similar, para la Sección B, la Tabla 2 muestra medias iniciales también equivalentes ($GE = 9.66$; $GC = 9.60$), sin evidencias de desigualdad en la prueba de contraste, esto asegura que las diferencias posteriores pueden atribuirse a la intervención y no a discrepancias previas.

7.2 Verificación de supuestos estadísticos y fiabilidad intercalificador

Antes de proceder con los análisis inferenciales, se verificaron los supuestos metodológicos fundamentales que garantizan la validez de las pruebas estadísticas aplicadas. En primer lugar, la normalidad de las puntuaciones de ganancia fue confirmada mediante la prueba de Shapiro-Wilk, la cual no evidenció desviaciones significativas en ninguno de los grupos, ni en la Sección A (GE: $p = .118$; GC: $p = .641$) ni en la Sección B (GE: $p = .562$; GC: $p = .776$). En segundo lugar, la

homogeneidad de varianzas se corroboró con la prueba de Levene, que indicó equivalencia en la dispersión de las puntuaciones entre los grupos tanto para la Sección A ($F = 0.115$, $p = .737$) como para la Sección B ($F = 2.058$, $p = .159$). Finalmente, se evaluó la ICC en la Sección B, observándose una consistencia calificada como buena a excelente. En el GE, la ICC pasó de .807 en el pretest a .901 en el posttest, mientras que en el GC pasó de .799 a .971. Estos resultados confirman la solidez de las mediciones de producción oral, respaldando tanto la confiabilidad del procedimiento como la pertinencia de los análisis estadísticos posteriores.

7.3 Análisis de los cambios intragrupo

7.3.1 Sección A: Reconocimiento auditivo

En la Sección A, los cambios observados fueron mínimos y no alcanzaron significación estadística en ninguno de los grupos. En el GE, la media se incrementó de 11.77 a 12.18, lo que representa una ganancia de apenas 0.41 puntos, sin diferencias significativas ($t(21) = 0.80$, $p = .430$) y con un tamaño del efecto muy pequeño ($d_z = 0.17$; IC 95 % Δ [-0.65, 1.47]). De forma similar, en el GC la media pasó de 12.05 a 12.24, con una ganancia de 0.19 puntos que tampoco resultó significativa ($t(20) = 0.39$, $p = .702$) y un tamaño del efecto prácticamente nulo ($d_z = 0.08$; IC 95 % Δ [-0.83, 1.21]). En conjunto, estos hallazgos indican que, durante el periodo de intervención, tanto la estrategia mediada por tecnología como la instrucción tradicional tuvieron un impacto nulo o muy limitado en el reconocimiento auditivo de pares mínimos.

Tabla 1 Resultados de la Sección A (Reconocimiento Auditivo)

Grupo	Pre M (DE)	Post M (DE)	Δ M (DE)	t(df)	p	d _z	IC95 % Δ
GE	11.77 (2.16)	12.18 (1.79)	0.41 (2.38)	0.80(21)	0.430	0.17	[-0.65, 1.47]
GC	12.05 (1.75)	12.24 (1.81)	0.19 (2.25)	0.39(20)	0.702	0.08	[-0.83, 1.21]

7.3.2 Sección B: Producción oral

En la Sección B, ambos grupos evidenciaron mejoras estadísticamente significativas y con una magnitud pedagógicamente relevante. En el GE, la puntuación media aumentó de 9.66 a 11.52, lo que representa una ganancia de 1.86 puntos altamente significativa ($t(21) = 3.51$, $p = .002$), acompañada de un tamaño del efecto grande ($dz = 0.75$; IC 95 % Δ [0.76, 2.97]). De manera paralela, el GC pasó de una media de 9.60 a 11.29, con una ganancia de 1.69 puntos igualmente significativa ($t(20) = 4.06$, $p = .001$) y un tamaño del efecto grande ($dz = 0.89$; IC 95 % Δ [0.82, 2.56]). Estos resultados indican que tanto la estrategia basada en GTM como la instrucción tradicional fueron altamente efectivas para potenciar la producción oral en un corto periodo de tiempo, confirmando que la habilidad de pronunciación es sensible a procesos de práctica intensiva independientemente del enfoque metodológico.

Tabla 2 Resultados de la Sección B (Producción oral)

Grupo	Pre M (DE)	Post M (DE)	Δ M (DE)	t(df)	p	dz	IC95 % Δ
GE	9.66 (3.05)	11.52 (3.42)	1.86 (2.49)	3.51(21)	0.002	0.75	[0.76, 2.97]
GC	9.60 (2.68)	11.29 (3.13)	1.69 (1.91)	4.06(20)	0.001	0.89	[0.82, 2.56]

Figura 2 Distribución de ganancias por grupo en la Sección A

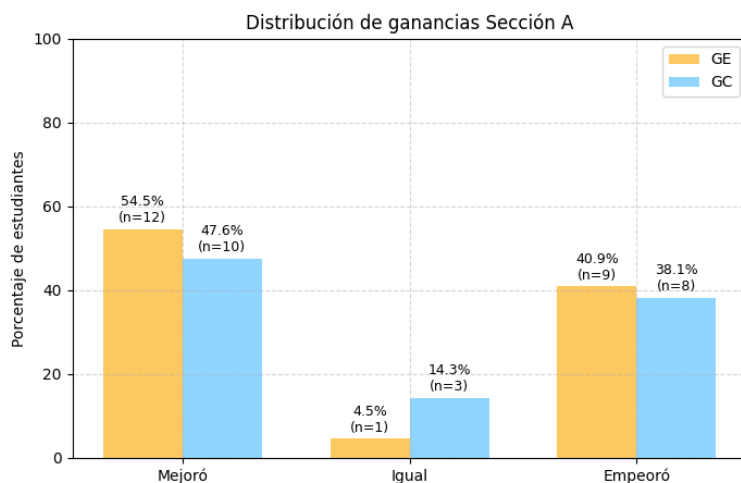
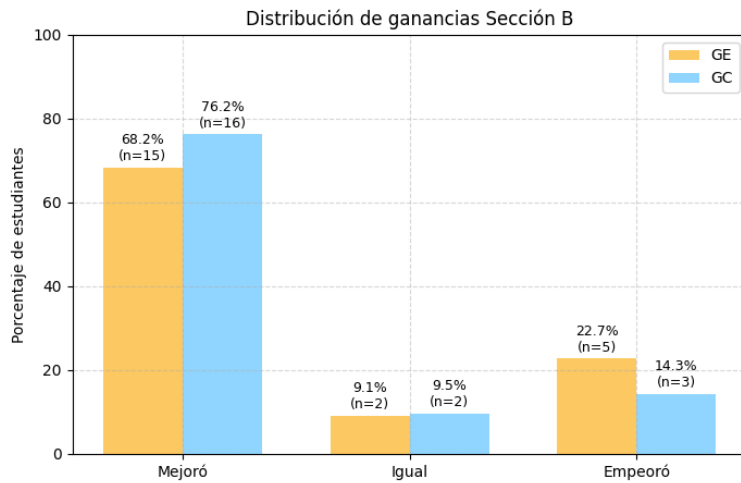


Figura 3 Distribución de ganancias por grupo en Sección B

7.4 Análisis de comparación entre grupos

En la comparación directa de las ganancias entre grupos, los resultados no evidenciaron ventajas significativas de un enfoque sobre el otro. En la Sección A, la diferencia en las ganancias fue mínima, de apenas 0.22 puntos a favor del GE, sin significación estadística ($t(41.0) = 0.31$, $p = .759$) y con un tamaño del efecto prácticamente nulo (Hedges' $g = 0.09$; IC 95 % [-0.49, 0.68]). En la Sección B, la diferencia favoreció al GE por 0.17 puntos, tampoco alcanzando significación ($t(39.2) = 0.26$, $p = .799$) y el tamaño del efecto fue pequeño (Hedges' $g = 0.08$; IC 95 % [-0.51, 0.66]). En síntesis, los hallazgos indican que, aunque ambos métodos de enseñanza lograron mejoras notables de manera independiente, sus efectos relativos fueron equivalentes durante el periodo de estudio.

7.5 Análisis de heterogeneidad y desempeño por ítems

El examen de la variabilidad individual y de los patrones fonológicos muestra que, aunque ambos grupos lograron avances globales, los resultados se distribuyeron de manera desigual entre estudiantes y categorías de sonidos. La Figura 2 muestra que, aunque predomina la categoría de mejora, la dispersión es elevada. En la Sección A, el 54.5 % de los estudiantes del GE mejoraron mientras que un 40.9 % empeoraron y en el GC el 47.6 % incrementó sus puntajes frente a un

38.1 % que descendió, lo que refleja diferencias en la sensibilidad individual a la instrucción recibida. Las categorías fonológicas en la Tabla 3 permiten matizar este panorama, ya que los avances fueron muy heterogéneos. En el GE, no se registró mejora en vocales (0.0 %), mientras que las oclusivas (4.4 %) y fricativas (3.0 %) mostraron incrementos leves en la Sección A, como se ilustra en la Figura 4. En el GC, las vocales apenas subieron un 0.5 %, las oclusivas incluso descendieron (-1.0 %) y las fricativas presentaron una mejora modesta (4.0 %). En conjunto, esto confirma que las ganancias en reconocimiento auditivo fueron limitadas y, en particular, que las vocales permanecieron como la categoría más débil en ambos grupos.

Tabla 3 Resultados por categoría fonológica en Sección A

Grupo	Categoría	Pre M (%)*	Post M (%)	Δ M (%)
GE	Vocales	6.00 (60.0)	6.00 (60.0)	0.00 (0.0)
	Oclusivas	2.55 (51.0)	2.77 (55.4)	0.22 (4.4)
	Fricativas	3.23 (53.8)	3.41 (56.8)	0.18 (3.0)
GC	Vocales	5.71 (57.1)	5.76 (57.6)	0.05 (0.5)
	Oclusivas	3.10 (62.0)	3.05 (61.0)	-0.05 (-1.0)
	Fricativas	3.24 (54.0)	3.48 (58.0)	0.24 (4.0)

Nota. Vocales (ítems 5, 6, 7, 8, 12, 17, 18, 19, 20); Oclusivas (ítems 1, 2, 3, 4, 16); Fricativas (ítems 9, 10, 11, 13, 14, 15). *Porcentaje de acierto calculado sobre el número de puntos máximos, igual al número de ítems por categoría.

En la Sección B la tendencia fue mucho más favorable. La Figura 3 permite visualizar como el 68.2 % de los estudiantes del GE y el 76.2 % del GC aumentaron sus puntajes, lo que evidencia una mayor receptividad de la producción a la práctica intensiva. Esta visualización refuerza la interpretación de eficacia general de las dos metodologías para la producción oral y contrasta con la dispersión observada en la Sección A. El análisis por categorías fonológicas en la Tabla 4 muestra que el GE obtuvo su mayor ganancia en vocales (15.7 %), seguida de mejoras relevantes en oclusivas (9.5 %) y más discretas en fricativas (4.1 %). Por su parte, el GC también alcanzó avances en las tres categorías, con incrementos similares en oclusivas (9.5 %) y fricativas (8.3 %), aunque más reducidos en vocales (7.5 %). Este patrón, visualizado mejor en la Figura 5,

sugiere que la mediación con GTM pudo haber favorecido de manera particular la producción de contrastes vocálicos, tradicionalmente más difíciles para los hispanohablantes, mientras que la instrucción tradicional tuvo un impacto más equilibrado en oclusivas y fricativas.

Tabla 4 Resultados por categoría fonológica en Sección B

Grupo	Categoría	Pre M (%)*	Post M (%)	ΔM (%)
GE	Vocales	2.36 (39.3)	3.30 (55.0)	0.94 (15.7)
	Oclusivas	3.75 (62.5)	4.32 (72.0)	0.57 (9.5)
	Fricativas	2.77 (46.2)	3.02 (50.3)	0.25 (4.1)
GC	Vocales	2.50 (41.7)	2.95 (49.2)	0.45 (7.5)
	Oclusivas	3.83 (63.8)	4.40 (73.3)	0.57 (9.5)
	Fricativas	2.62 (43.7)	3.12 (52.0)	0.50 (8.3)

Nota. Vocales (ítems 2, 5, 8); Oclusivas (ítems 1, 7, 10); Fricativas (ítems 3, 4, 9). *Porcentaje de acierto calculado sobre el número de puntos máximos, igual a 6.

Figura 4 Mejora porcentual por categoría fonológica en Sección A

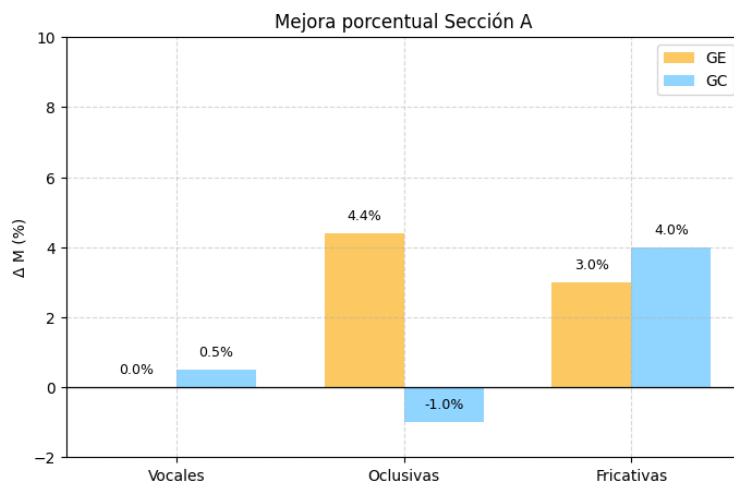
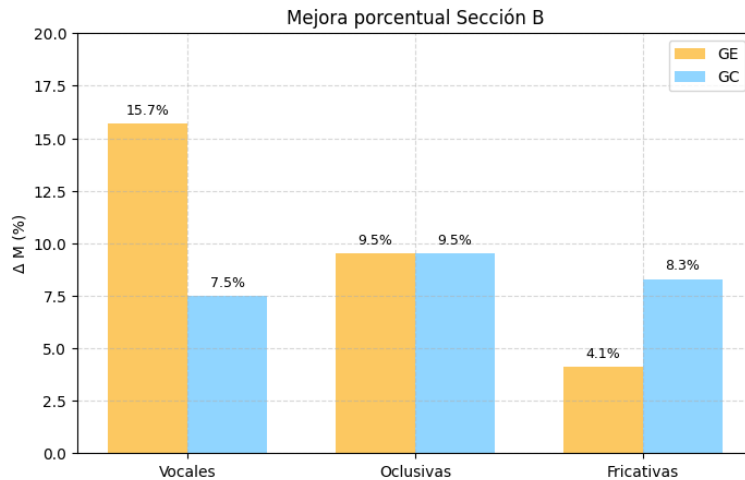


Figura 5 Mejora porcentual por categoría fonológica en Sección B

7.6 Encuesta de percepción del GE

La encuesta aplicada al GE al finalizar la intervención revela percepciones mayoritariamente positivas hacia la experiencia constructora con GTM, alcanzando medias generales en todas las dimensiones superiores a 4, ver Tabla 3. Este panorama se ve reforzado por los comentarios abiertos de los estudiantes y por las observaciones sistemáticas del docente, que en conjunto describen un clima socioafectivo favorable, con alta participación, colaboración y disfrute sostenido.

7.6.1 Dimensión 1: Motivación y disfrute

- “Me divertí con las actividades”: $M = 4.81$ ($DE = 0.40$).
- “Me gustó trabajar en equipo”: $M = 4.56$ ($DE = 0.92$).

Los comentarios cualitativos corroboran estas valoraciones con expresiones como “me encantó trabajar en equipo” y “fue muy divertido”, lo que sugiere que la novedad tecnológica y el carácter práctico de la construcción de modelos favorecieron la implicación emocional y el interés sostenido. La bitácora docente coincide plenamente: se registró 100 % de participación,

colaboración y enfoque en las sesiones observadas, indicando que la motivación no fue episódica sino estable a lo largo del proceso.

7.6.2 Dimensión 2: Autoeficacia y confianza

- “Pude realizar la actividad propuesta”: $M = 4.61$ ($DE = 0.49$).
- “El modelo me ayudó a escuchar mejor mi pronunciación”: $M = 4.58$ ($DE = 0.67$).

En los comentarios emergen alusiones a logro y dominio de la tarea (“pude entender más”, “me ayudó a pronunciar mejor”), que se alinean con el aumento de la autopercepción de capacidad y con la disminución de la frustración hacia el cierre de la intervención, ver Figura 6. Este patrón sugiere un andamiaje efectivo, ya que los retos iniciales dieron paso a una mayor confianza conforme avanzó la familiaridad con la herramienta y se consolidaron rutinas de práctica.

7.6.3 Dimensión 3: Usabilidad y utilidad percibida

- “Quiero seguir practicando con esta estrategia”: $M = 4.47$ ($DE = 0.53$).

El deseo de continuidad reportado por los estudiantes se complementa con descripciones cualitativas de facilidad de uso con acompañamiento docente y utilidad directa para la pronunciación (“me ayudó a escuchar mejor mi pronunciación”). En la bitácora, el 100 % de iniciativa e interés por explorar funcionalidades respalda la percepción de usabilidad y valor práctico, elementos clave para la sostenibilidad de la práctica más allá del aula.

7.6.4 Dimensión 4: Experiencia constructorista

Los comentarios abiertos evidencian entusiasmo y satisfacción (“me encantó trabajar en equipo”, “fue muy divertido”, “me ayudó a escuchar mejor mi pronunciación”), junto con observaciones puntuales de estudiantes que prefirieron trabajar de forma individual hacia el final. Esta preferencia emergente por la autonomía es coherente con el incremento de la autoeficacia descrito arriba: a medida que los estudiantes dominaron la lógica de entrenamiento y prueba de

modelos, algunos pasaron de valorar el soporte del equipo a buscar itinerarios de práctica individual para profundizar en sus propias dificultades fonéticas.

Tabla 5 Resultados de la encuesta de percepción en GE

Dimensión	Ítems	M general	Rango
Motivación y disfrute	2	4.69	3–5
Autoeficiencia y confianza	2	4.60	2–5
Usabilidad y utilidad	1	4.47	3–5

7.7. Observaciones sistemáticas del docente investigador

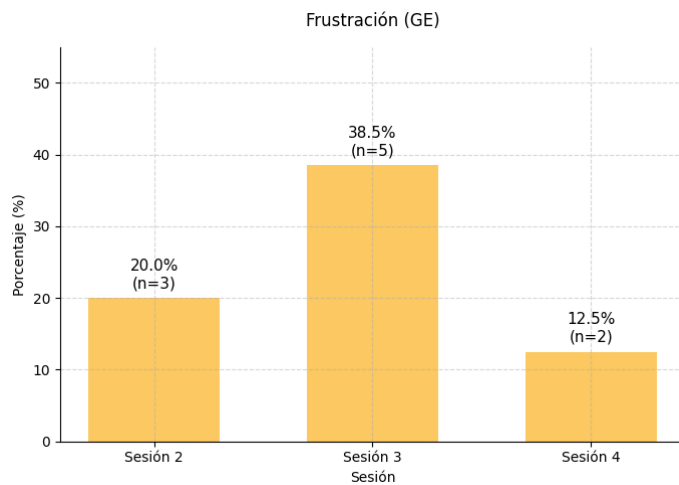
Las observaciones estructuradas documentan comportamientos y manifestaciones socioafectivas durante cuatro sesiones (la primera introductoria y general) en el GE, con un total de 44 registros individuales de participación. En los indicadores de compromiso académico, la participación, persistencia ante retos e iniciativa para explorar funcionalidades se registran en el 100 % de las observaciones; en los indicadores socioafectivos, se evidencia colaboración efectiva y expresiones de satisfacción en la totalidad de los casos. En contraste, las manifestaciones de frustración se observan en un 22.7 % de los registros, con una evolución temporal característica: aumento en la sesión 3 y descenso en la sesión 4.

Esta curva de frustración se interpreta cualitativamente como un reto óptimo. La sesión 3, más demandante por la carga técnica y el ajuste fino de categorías, tensionó la actividad; sin embargo, el acompañamiento docente y la progresiva competencia de los estudiantes permitieron convertir el desafío en aprendizaje, reduciendo la frustración en la sesión 4 y consolidando la confianza y el desempeño.

En conjunto, las voces de los estudiantes y la mirada del docente delinear una experiencia atractiva, útil y formativa, en la que la motivación intrínseca, la autoeficacia y la colaboración actúan como palancas del aprendizaje. La preferencia emergente por la práctica individual sugiere oportunidades de personalización, especialmente en contrastes vocálicos complejos, y de práctica

distribuida fuera del aula, mientras que la gestión del reto observada a lo largo de las sesiones constituye un mecanismo clave para transformar la dificultad en progreso consolidado.

Figura 6 Indicador de frustración por sesión en el GE



Nota. Los porcentajes corresponden al total de estudiantes observados en cada sesión.

8 Discusión

El propósito de este capítulo es realizar una interpretación crítica y un análisis profundo de los resultados obtenidos, contextualizándolos dentro del marco de la literatura académica existente. Se busca establecer un diálogo entre los hallazgos de esta investigación y los estudios previos, con el fin de comprender las implicaciones, contradicciones y contribuciones de utilizar una propuesta didáctica constructorista con GTM para la enseñanza de la pronunciación en inglés.

8.1 Efectividad de la intervención y la ausencia de diferencias significativas

El primer hallazgo central de esta investigación es que tanto la intervención didáctica con GTM como la instrucción fonética tradicional generaron mejoras estadísticamente significativas en la competencia fonética de los estudiantes, especialmente en la producción oral (Sección B). Este resultado es consistente con una amplia base de evidencia que confirma la eficacia de los sistemas CAPT en diversos contextos educativos. Estudios como los de Mak et al. (2003), Neri et al. (2008) y más recientemente, López-Minotta et al. (2025) han demostrado que la retroalimentación automatizada produce mejoras cuantificables en la pronunciación. Por lo tanto, el progreso del GE valida el potencial de las herramientas de IA, como GTM, para complementar la enseñanza fonética.

Sin embargo, el hallazgo más provocador es la ausencia de una diferencia estadísticamente significativa en la magnitud de la mejora entre el GE y el de control. Este resultado contrasta con estudios como el de Tejedor-García et al. (2020), quienes reportaron una mejora relativa del 205 % en el grupo que utilizó CAPT frente a un 82 % en el grupo con instrucción tradicional. Varias hipótesis pueden explicar esta discrepancia: en primer lugar, la corta duración de la intervención (cuatro sesiones) pudo ser insuficiente para que las ventajas del enfoque con GTM se tradujeran en una brecha estadísticamente mayor. En segundo lugar, la efectividad del docente en la aplicación de la metodología tradicional en el grupo de control no debe ser subestimada; como lo sugiere la mejora significativa alcanzada en este grupo, una instrucción fonética bien ejecutada sigue siendo pedagógicamente potente.

Finalmente, como señalan Pujari et al. (2022), la precisión de sistemas como GTM depende de grandes volúmenes de datos; en este caso, los modelos creados por los estudiantes con muestras

limitadas de audio pudieron no haber ofrecido una retroalimentación tan precisa como la de un docente experimentado. No obstante, un análisis más detallado por categorías fonológicas (ver sección 8.3) reveló patrones diferenciados que enriquecen la interpretación de los resultados globales.

8.2 Construccionismo: motivación y aprendizaje activo

Si bien los resultados cuantitativos no mostraron superioridad, es en el análisis cualitativo donde reside la contribución más significativa de este estudio. La intervención valida el fundamento teórico del construccionismo, al demostrar que el aprendizaje es más significativo cuando los estudiantes están activamente construyendo artefactos de aprendizaje. En este caso, los estudiantes no fueron meros usuarios de un software, sino programadores y creadores de sus propios modelos de clasificación. Este giro metodológico, que coincide con lo señalado por Carney et al. (2020) y Sayginer y Tekdal (2023) sobre el potencial de GTM, explica el elevado nivel de compromiso observado.

La experiencia del GE trascendió la simple práctica fonética y se vinculó con el desarrollo de la alfabetización digital y el pensamiento computacional. Tal como sugiere Dwivedi et al. (2024), la llamada “curva de aprendizaje de la frustración,” un pico de dificultad seguido por resolución exitosa constituye evidencia de aprendizaje profundo a través del ensayo y error, pilar del enfoque construccionista.

Asimismo, la mecánica de GTM ofrece un ejemplo práctico de la Hipótesis del *Noticing* de Schmidt. Para que un estudiante lograra que el modelo clasificara correctamente su pronunciación, debía primero notar la discrepancia entre su producción y el sonido objetivo. Como señalan Ashwell y Elam (2017), los fallos de un sistema ASR funcionan como retroalimentación externa y objetiva que obliga al aprendiz a reformular su enunciado. Esta retroalimentación inmediata, desprovista del juicio social que puede ocurrir en la interacción humana, genera un entorno de baja ansiedad, ideal para la práctica de la pronunciación (Neri, 2007; Tejedor-García, 2020).

8.3 Resultados por categorías fonológicas

El análisis desagregado por categorías fonológicas aporta matices adicionales a la comprensión de la efectividad de la intervención. Ambos grupos evidenciaron mejoras estadísticamente significativas en la producción oral de pares mínimos, confirmando que la práctica fonética intensiva genera efectos positivos independientemente de la mediación tecnológica. Sin embargo, los patrones no fueron uniformes.

En el GE, las vocales mostraron las mayores ganancias, alcanzando significancia estadística en contraste con el GC, que no logró progresos relevantes en esta categoría. Este hallazgo sugiere que la retroalimentación visual inmediata de GTM resulta particularmente eficaz para los contrastes de duración vocálica, perceptualmente más salientes para hispanohablantes. En contraste, tanto el GE como el GC mejoraron en oclusivas, lo que indica que estos sonidos responden de manera consistente a la retroalimentación, ya sea automatizada o docente. Finalmente, en el caso de las fricativas, solo el GC alcanzó mejoras significativas, lo cual apunta a que ciertos contrastes consonánticos requieren apoyos adicionales de tipo articulatorio o intervenciones más prolongadas para consolidarse.

8.4 Dimensión afectivo-motivacional y comparación con la literatura

La ausencia de diferencias significativas entre grupos en los puntajes globales no debe opacar los hallazgos en la dimensión afectivo-motivacional, donde la propuesta con GTM generó un valor agregado evidente. Los altos niveles de participación activa, colaboración entre pares y satisfacción expresados por los estudiantes, recogidos en encuestas de percepción y observaciones del docente, corroboran que GTM propició un entorno de aprendizaje dinámico, retador y motivador. La autonomía se vio favorecida por el hecho de que los estudiantes crearan y entrenaran sus propios modelos, lo que les permitió monitorear su progreso y consolidar progresivamente su autopercepción de competencia. La mencionada “curva de la frustración” funcionó aquí como catalizador de resiliencia y autoeficacia, componentes centrales del aprendizaje constructivista.

Estos hallazgos se alinean con investigaciones previas sobre CAPT. Mak et al. (2003) mostraron que, si bien el 73 % de los estudiantes se benefició significativamente de sistemas ASR, un 27 % respondió mejor a la instrucción tradicional, sugiriendo que la efectividad no es universal,

sino dependiente de factores contextuales e individuales. Neri et al. (2008) observaron avances proporcionales a los de este estudio en niños de 11 años, lo cual refuerza su pertinencia metodológica. En contraste, Tejedor-García et al. (2020) y López-Minotta et al. (2025) reportaron mejoras de mayor magnitud, explicables por intervenciones más largas, estudiantes de mayor madurez cognitiva y el uso de sistemas CAPT comerciales más sofisticados.

En este sentido, el aporte de la presente investigación no radica en replicar incrementos absolutos de otros contextos, sino en demostrar que una herramienta gratuita y accesible como GTM puede generar efectos comparables a los de metodologías tradicionales y, adicionalmente, potenciar dimensiones afectivas que los modelos convencionales difícilmente abordan. Esto refuerza la visión del construccionismo digital como marco pedagógico capaz de integrar lo cognitivo y lo motivacional, aportando evidencia empírica desde un contexto de educación primaria latinoamericano.

9 Conclusiones

Este capítulo presenta las interpretaciones finales de la investigación, sintetizando los hallazgos cuantitativos y cualitativos para dar respuesta a los objetivos general y específicos que guiaron el estudio. Las conclusiones se derivan directamente del análisis de los datos y confirman o refutan las hipótesis formuladas en el planteamiento del problema. La presente investigación tuvo como propósito central evaluar el efecto de una propuesta didáctica innovadora, basada en el construccionismo y el aprendizaje activo con GTM, sobre el desarrollo de la competencia fonética en estudiantes de quinto grado. A continuación, se exponen las conclusiones articuladas con cada uno de los objetivos.

El objetivo general buscaba evaluar el efecto de la propuesta didáctica con GTM sobre el reconocimiento auditivo y la producción oral de pares mínimos en inglés, comparando los resultados del GE con un grupo de control. Se concluye que la intervención didáctica basada en GTM tuvo un efecto positivo y estadísticamente significativo en la mejora de la competencia fonética de los estudiantes. Ambos grupos, tanto el experimental como el de control, mostraron mejoras significativas en sus puntuaciones totales entre el pretest y el postest, particularmente en la dimensión de producción oral. Sin embargo, al comparar la efectividad relativa entre ambas modalidades, no se encontraron diferencias estadísticamente significativas ($p < .05$) en la magnitud de la mejora entre el GE y el grupo de control.

En contraste, los resultados en reconocimiento auditivo (Sección A) fueron limitados, sin mejoras estadísticamente significativas en ninguno de los grupos (GE: 0.41; GC: 0.19, y con una alta variabilidad individual en los puntajes. Una proporción considerable de estudiantes mantuvo o incluso disminuyó sus resultados (40.9 % en GE y 38.1 % en GC), lo que sugiere que los contrastes perceptivos requieren intervenciones más prolongadas o apoyos adicionales para consolidarse. Este hallazgo complementa la visión de que la producción oral es más sensible a la práctica intensiva en periodos cortos, mientras que la percepción fonética demanda un andamiaje sostenido.

En lo que respecta al primer objetivo específico, la propuesta didáctica fue diseñada e implementada exitosamente conforme a los principios del construccionismo. Durante las cuatro sesiones, los estudiantes trascendieron el rol de receptores pasivos para convertirse en creadores

activos de sus propias herramientas de aprendizaje, validando la viabilidad del enfoque propuesto en un contexto escolar real.

En relación con el segundo objetivo, los datos cuantitativos validaron un progreso tangible y estadísticamente significativo en el GE. El análisis pretest – postest confirmó que la interacción con la herramienta se tradujo en una mejora medible, particularmente en la producción oral, lo que demuestra la pertinencia de la intervención para el desarrollo de habilidades fonéticas prácticas.

El tercer objetivo, que comparaba el desempeño entre grupos, arrojó uno de los resultados más reveladores. No se encontraron diferencias estadísticamente significativas que demostraran la superioridad de la intervención con GTM sobre el método tradicional. La heterogeneidad en las respuestas refuerza esta conclusión: mientras que en producción oral la mayoría de estudiantes mejoró (68.2 % en GE y 76.2 % en GC), en reconocimiento auditivo una parte considerable no logró avances. Este hallazgo subraya la robustez de las prácticas pedagógicas establecidas y sugiere que la innovación tecnológica debe ser vista como un complemento potente, más que como un reemplazo.

En cuanto al cuarto objetivo, el análisis de la percepción estudiantil arrojó los resultados más contundentes y positivos. La experiencia de los participantes del GE fue excepcionalmente favorable, con niveles máximos de motivación, compromiso y disfrute. Las encuestas revelaron medias superiores a 4.5 en todas las dimensiones (motivación, confianza, usabilidad), mientras que las observaciones sistemáticas documentaron 100 % de participación, colaboración y persistencia ante las dificultades. La “curva de frustración” —con un incremento en la tercera sesión seguido de un descenso en la cuarta— evidenció que los retos pedagógicos fueron gestionados con éxito, consolidando resiliencia y autoeficacia.

Un análisis más profundo de los resultados por categorías fonológicas revela que la intervención tuvo un impacto diferenciado según el tipo de contraste. El GE se destacó en la mejora de vocales (+15.7 %), mientras que el GC mostró un desempeño más equilibrado en oclusivas (+9.5 %) y fricativas (+8.3 %). De manera particular, la ganancia significativa en fricativas interdentales (/θ–ð/) constituye un hallazgo relevante, ya que corresponde a uno de los contrastes más difíciles para los hispanohablantes. Este patrón sugiere que la retroalimentación inmediata de GTM favoreció especialmente los contrastes vocálicos, mientras que la instrucción tradicional proporcionó apoyos más consistentes en consonantes.

De este modo, la investigación confirma que el enfoque constructorista —donde el estudiante se convierte en creador de tecnología— genera un entorno de aprendizaje positivo que difícilmente puede ser replicado por métodos tradicionales. Finalmente, se confirma la hipótesis de trabajo en cuanto a que la propuesta con GTM genera una mejora significativa en el aprendizaje, pero se rechaza la parte que anticipaba una superioridad estadística frente al grupo de control. La hipótesis nula, que postulaba la ausencia de diferencias significativas entre los grupos, no pudo ser rechazada.

En definitiva, el valor de la innovación no reside únicamente en generar ventajas cuantitativas inmediatas, sino en su capacidad para transformar la experiencia de aprendizaje: fomentar la motivación y la resiliencia, desarrollar competencias digitales y preparar a los estudiantes para los desafíos de la educación contemporánea.

10 Recomendaciones

Derivado de un análisis crítico de los resultados, las conclusiones y las limitaciones inherentes a este estudio, se formulan recomendaciones orientadas a refinar la propuesta pedagógica y a guiar futuras investigaciones en el campo de la tecnología educativa y la enseñanza de lenguas. Reconocer las limitaciones y los desafíos encontrados permite delinear una hoja de ruta más sólida para quienes deseen replicar o ampliar la experiencia.

10.1 Recomendaciones para la práctica pedagógica

10.1.1 Brevidad de la intervención

El principal factor que pudo limitar el impacto cuantitativo de la propuesta fue su corta duración de cuatro sesiones. Este tiempo fue insuficiente para que las ventajas motivacionales del enfoque con GTM se tradujeran en una diferencia de rendimiento estadísticamente superior a la del grupo de control. Futuras implementaciones deberían extenderse a lo largo de un período académico más prolongado, como un bimestre o un semestre, de modo que el aprendizaje se consolide, las habilidades maduren y se puedan medir cambios más estables en el desempeño lingüístico.

10.1.2 Dependencia tecnológica y desafíos logísticos

La bitácora del proyecto evidenció que fallas en los equipos y limitaciones de tiempo impidieron que algunos estudiantes completaran las actividades según lo planeado. Para reducir esta dependencia, se recomienda diseñar planes de contingencia, como actividades “desconectadas” que refuercen los mismos conceptos, o guías de trabajo en casa más robustas que permitan continuar la práctica incluso en ausencia de recursos tecnológicos. Asimismo, la institución debería considerar el fortalecimiento de su infraestructura y la capacitación técnica de los docentes para evitar interrupciones en el proceso formativo.

10.1.3 Precisión de los modelos creados por estudiantes

A diferencia de los sistemas CAPT comerciales, la precisión de los modelos en GTM depende directamente de la calidad y cantidad de los datos aportados por los estudiantes. Un posible fallo fue que los modelos, entrenados con pocas muestras de voz infantil, no alcanzaran un nivel de precisión óptimo para ofrecer retroalimentación confiable. Como mejora, se sugiere que el docente guíe un proceso más riguroso de “curación de datos”, orientando a los estudiantes en la grabación de audios más limpios, el incremento del número de muestras y la validación cruzada de resultados. Esto no solo elevaría la calidad técnica, sino que también fortalecería competencias en alfabetización digital y pensamiento computacional.

10.1.4 Adoptar un modelo pedagógico híbrido y flexible

En lugar de plantear una dicotomía entre tecnología y tradición, se recomienda a los docentes integrar GTM como un componente dentro de un ecosistema de aprendizaje más amplio. La instrucción directa puede cubrir la teoría fonética, mientras que GTM potencia la práctica activa, lúdica y personalizada. Esta combinación híbrida y flexible equilibra el rigor académico con la motivación estudiantil, alineándose con las demandas de un enfoque de educación digital adaptativa.

10.1.5 Diseñar una transición del trabajo colaborativo al individual

Reconociendo el deseo de autonomía de los estudiantes, es pertinente plantear una implementación por fases. Se sugiere iniciar con trabajo en equipo, lo que facilita superar la barrera tecnológica inicial y fomenta la colaboración, para luego avanzar hacia proyectos individuales en los que cada estudiante pueda profundizar a su propio ritmo y asumir plena autoría de su creación. Este tránsito progresivo responde tanto a necesidades pedagógicas como al desarrollo de competencias para el aprendizaje autónomo.

10.2 Recomendaciones para futuras investigaciones

10.2.1 Estudios longitudinales sobre el impacto acumulativo

Para superar la limitación de la corta duración, es crucial realizar investigaciones a largo plazo. Estudios longitudinales permitirían indagar si el alto compromiso que genera el enfoque constructorista se traduce, con el tiempo, en una ventaja cuantitativa significativa, en una mayor retención de conocimientos y en la transferencia de habilidades a otros contextos comunicativos.

10.2.2 Diseños experimentales con aleatorización

La naturaleza cuasiexperimental de este estudio, con grupos preexistentes, constituye una limitación metodológica. Futuras investigaciones ganarían en validez interna si implementan un diseño experimental puro con asignación aleatoria de participantes a los grupos de control y experimental. De esta forma, se podría aislar con mayor certeza el efecto de la intervención, minimizando sesgos de selección.

10.2.3 Estudios comparativos de herramientas de IA

Es de gran interés académico realizar investigaciones que comparen la efectividad del enfoque constructorista de GTM —donde el alumno participa en la creación del modelo— con aplicaciones CAPT tradicionales y preentrenadas, como ELSA Speak. Este tipo de estudios ayudaría a discernir el peso relativo de la precisión técnica de la retroalimentación frente al valor pedagógico del aprendizaje activo y la agencia estudiantil.

10.2.4 Investigación sobre variables afectivas y metacognitivas

Los hallazgos cualitativos de este trabajo constituyen un aporte relevante. Futuras investigaciones deberían medir cuantitativamente, mediante escalas validadas, el impacto de este tipo de intervenciones sobre variables como autoeficacia, motivación intrínseca, ansiedad al hablar y autonomía en el aprendizaje. Esto permitiría cuantificar los beneficios observados

cualitativamente y ampliar la comprensión de los factores emocionales y metacognitivos que median en el éxito de la pronunciación en L2.

En conjunto, estas recomendaciones, surgidas de una reflexión autocrítica sobre las limitaciones y hallazgos del estudio, trazan una ruta clara para avanzar en la enseñanza de lenguas mediada por inteligencia artificial accesible. El camino futuro no consiste en elegir entre innovación tecnológica y práctica tradicional, sino en construir una síntesis pedagógica adaptativa que maximice el potencial de ambas. Al atender las mejoras sugeridas y explorar las líneas de investigación planteadas, la comunidad educativa podrá avanzar hacia una comprensión más completa del verdadero potencial de herramientas como GTM, no solo como optimizadoras del rendimiento, sino como catalizadoras de un aprendizaje más autónomo, motivador y significativo.

Referencias

- Acosta Acosta, E. L. (2023). *Minimal pair technique and English pronunciation* [Tesis de licenciatura, Universidad Técnica de Ambato]. Repositorio de la Universidad Técnica de Ambato. <https://repositorio.uta.edu.ec/handle/123456789/39004>
- Aryanti, R., & Santosa, M. (2024). A systematic review on artificial intelligence applications for enhancing EFL students' pronunciation skill. *The Art of Teaching English as a Foreign Language*, 5(1), 102–113. <https://doi.org/10.36663/tatefl.v5i1.718>
- Ashwell, T., & Elam, J. (2017). How accurately can the Google Web Speech API recognize and transcribe Japanese L2 English learners' oral production? *JALT CALL Journal*, 13(1), 59–76. <https://doi.org/10.29140/jaltcall.v13n1.j212>
- Asmaul Husna, S., Sholikhah, S., & Lubis, Y. (2023). An analysis of minimal pairs of consonant sounds in English. *AGTMosfer: Jurnal Pendidikan, Bahasa, Sastra, Seni, Budaya, Dan Sosial Humaniora*, 1(3), 92–97. <https://doi.org/10.59024/aGTMosfer.v1i3.208>
- Badilla Saxe, E., & Chacón Murillo, A. (2004). Construcciónismo: Objetos para pensar, entidades públicas y micromundos. *Revista Electrónica Actualidades Investigativas en Educación*, 4(1), 1–20. Disponible en: <https://www.redalyc.org/articulo.oa?id=44740104>
- Báez, C., Buchner, J., Ullrich, A.-L., & Schallert-Vallaster, S. (2023). Exploring the perception of AI in learning: Unveiling the role of student and teacher motivation and self-efficacy. *OSF Preprints*. <https://doi.org/10.31219/osf.io/wqvrh>
- Bejtíc, Z. (2023). Corrective vs. nurturing feedback in design education: Alternative models of critique that positively impact students' sense of self-efficacy. En D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Creative approaches to technology-enhanced learning for the workplace and higher education* (pp. 39–54). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-031-41637-8_4
- Benavides, J. E. (2021). Level of English in Colombian higher education: A decade of stagnation. *Profile: Issues in Teachers' Professional Development*, 23(1), 57–73. <https://doi.org/10.15446/profile.v23n1.83135>
- Bu, Y., Ma, T., Li, W., Zhou, H., Jia, J., Chen, S., Xu, K., Shi, D., Wu, H., Yang, Z., Li, K., Wu, Z., Shi, Y., Lu, X., & Liu, Z. (2021). PTeacher: A computer aided personalized pronunciation training system with exaggerated audio visual corrective feedback (arXiv:2105.05182). *arXiv*. <https://arxiv.org/abs/2105.05182>
- Carlet, A. (2017). *L2 perception and production of English consonants and vowels by Catalan speakers: The effects of attention and training task in a cross-training study* [Tesis doctoral, Universitat Autònoma de Barcelona]. Universitat Autònoma de Barcelona.

- Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., Jongejan, J., Pitaru, A., & Chen, A. (2020). Teachable Machine: Approachable web-based tool for exploring machine learning classification. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1–8). Association for Computing Machinery. <https://doi.org/10.1145/3334480.3382839>
- Chandran, M., Wang, S. L., Dzulkifly, S., Ng, T. F., & Ghoneim, A. S. (2024). Enhancing English pronunciation learning in primary education through natural language processing: A quantitative study. *Journal of ICT in Education*, 11(2), 1–17. <https://doi.org/10.37134/jictie.vol11.2.1.2024>
- Duisenova, M. M., & Zhorabekova, A. N. (2024). The efficacy of gamification and artificial intelligence in enhancing the motivation and efficacy of primary school kids in learning English. *PEDS*, 73(2), Article 027. <https://doi.org/10.48371/PEDS.2024.73.2.027>
- Dwivedi, U., Elsayed-Ali, S., Bonsignore, E., & Kacorri, H. (2024). Exploring AI problem formulation with children via Teachable Machines. En *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Article 273, 18 pp.). Association for Computing Machinery. <https://doi.org/10.1145/3613904.3642692>
- Forero-Corba, W., & Negre Bennasar, F. (2024). Técnicas y aplicaciones del Machine Learning e inteligencia artificial en educación: una revisión sistemática. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 209–253. <https://doi.org/10.5944/ried.27.1.37491>
- Gerson, M. (2006). Local and dialogic language ability and its implication for language teaching and testing. In R. Hughes (Ed.), *Spoken English, TESOL and applied linguistics: Challenges for theory and practice* (pp. 271–286). Palgrave Macmillan. https://doi.org/10.1057/9780230584587_12
- Guerrero-Nieto, C. H. (2024). English—A global South perspective: The case of a national bilingualism program in Colombia. En C. Leung & J. Lewkowicz (Eds.), *The Routledge companion to English studies* (2nd ed., pp. 269–282). Routledge. <https://doi.org/10.4324/9781003221265-24>
- Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, 126(2), 866–877. <https://doi.org/10.1121/1.3148196>
- Koneru, G. K. (2023). Ensuring optimal performance in online learning of STEM subjects: An autoethnographic study. En D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Creative approaches to technology-enhanced learning for the workplace and higher education* (pp. 301–313). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-031-41637-8_24

- Lai, K. W. K., & Chen, H. J. H. (2022). An exploratory study on the accuracy of three speech recognition software programs for young Taiwanese EFL learners. *Interactive Learning Environments*, 32(5), 1582–1596. <https://doi.org/10.1080/10494820.2022.2122511>
- Leow, R. (2019). Noticing hypothesis. In J. I. Lontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–7). Wiley. <https://doi.org/10.1002/9781118784235.eelt0086.pub2>
- López-Minotta, K. L., Chiappe, A., & Mella-Norambuena, J. (2025). Implementation of artificial intelligence to improve English oral expression. *Multidisciplinary Journal of Educational Research*, 15(1), 43–71. <http://dx.doi.org/10.17583/remie.16188>
- Mak, B., Siu, M., Ng, M., Tam, Y.-C., Chan, Y.-C., Chan, K.-W., Leung, K.-Y., Ho, S., Chong, F.-H., Wong, J., & Lo, J. (2003). PLASER: Pronunciation learning via automatic speech recognition. En *Proceedings of the HLT-NAACL 03 Workshop on Building Educational Applications Using Natural Language Processing* (Vol. 2, pp. 23–29). Association for Computational Linguistics. <https://doi.org/10.3115/1118894.1118898>
- Mao, W., & Wang, F.-Y. (2012). Forecasting group behavior via probabilistic plan inference. In W. Mao & F.-Y. Wang (Eds.), *New advances in intelligence and security informatics* (pp. 33–44). Academic Press. <https://doi.org/10.1016/B978-0-12-397200-2.00004-X>
- Maulana, A., & Suharto, R. (2024). Exploring the impact of automatic speech recognition on students' speaking skills and perceptions. *Jurnal Ilmu Komputer dan Teknologi*, 5(3), 48–55. <https://doi.org/10.35960/ikomti.v5i3.1661>
- Mora, J. C., Rochdi, Y., & Kivistö de Souza, H. (2014). Mimicking accented speech as L2 phonological awareness. *Language Awareness*, 23(1–2), 57–75. <https://doi.org/10.1080/09658416.2013.863898>
- Neri, A. (2007). *The pedagogical effectiveness of ASR-based computer assisted pronunciation training* [Tesis doctoral, Radboud Universiteit Nijmegen]. Radboud Universiteit Nijmegen. <https://hdl.handle.net/2066/44830>
- Neri, A., Cucchiaroni, C., & Strik, H. (2002). Feedback in computer assisted pronunciation training: Technology push or demand pull? En *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP 2002)* (pp. 1209–1212). Association for Computational Linguistics; ISCA. <https://doi.org/10.21437/ICSLP.2002-246>
- Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393–408. <https://doi.org/10.1080/09588220802447651>
- Okagbue, E. F., Ezeachikulo, U. P., Akintunde, T. Y., Tsakuwa, M. B., Ilokanulo, S. N., Obiasoanya, K. M., Ilodibe, C. E., & Ouattara, C. A. T. (2023). A comprehensive overview of artificial intelligence and machine learning in education pedagogy: 21 years

- (2000–2021) of research indexed in the Scopus database. *Social Sciences & Humanities Open*, 8(1), Article 100655. <https://doi.org/10.1016/j.ssaho.2023.100655>
- Ortega-Ochoa, E. (2022). A pedagogical conversational agent for tutoring in the development of educational research projects. En D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Innovative approaches to technology-enhanced learning for the workplace and higher education* (pp. 544–556). Springer. https://doi.org/10.1007/978-3-031-21569-8_51
- Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives*. Palgrave Macmillan. <https://doi.org/10.1057/978-1-137-47677-7>
- Pujari, Y. P., Prasad, D., Malleswari, N., & Gupta, N. (2022). Implementation of machine learning based Google Teachable Machine in early childhood education. *International Journal of Early Childhood Special Education*, 14, 4132–4138.
- Rizvi, S., Waite, J., & Sentance, S. (2023). Artificial intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review. *Computers and Education: Artificial Intelligence*, 4, 100145. <https://doi.org/10.1016/j.caeai.2023.100145>
- Sabaruddin, S., Purnama, A., Harmilawati, H., Aminah, S., Irmayani, I., Takdir, T., & AR, A. (2024). Artificial intelligence in English language teaching: A study of English teachers' TPACK. *Globish: An English-Indonesian Journal for English, Education, and Culture*, 13(2), 125–136. <https://doi.org/10.31000/globish.v13i2.11250>
- Saygıner, Ş., & Tekdal, M. (2023). The potential of using the Teachable Machine application for machine learning in education. En *Proceedings of the 16th International Computer and Instructional Technologies Symposium* (pp. 14–22). Eskişehir, Turkey.
- Senowarsito, S., & Ardini, S. (2023). The use of artificial intelligence to promote autonomous pronunciation learning: Segmental and suprasegmental features perspective. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 8(2), 133–150. <https://doi.org/10.21093/ijeltal.v8i2.1452>
- Shehata, M. (2024). A program based on artificial intelligence to enhance prospective teachers' English pronunciation. *CDELTA Occasional Papers in the Development of English Education*, 86, 145–179. <https://doi.org/10.21608/opde.2024.362818>
- Silva, M., Saraiva, A., Malta, D., Silva, J., Silva, R., & Santos, S. (2024). The integration of artificial intelligence into the personalization of education: A new paradigm for basic education. *ARACÊ*, 6(3). <https://doi.org/10.56238/arev6n3-100>
- Silva Valencia, J. C. (2022). A comparative linguistic analysis of English and Spanish phonological system. *GIST – Education and Learning Research Journal*, 25, 140–156. <https://doi.org/10.26817/16925777.1152>

- Su, J., Ng, D. T. K., & Chu, S. K. W. (2023). Artificial intelligence (AI) literacy in early childhood education: The challenges and opportunities. *Computers and Education: Artificial Intelligence*, 4, 100124. <https://doi.org/10.1016/j.caeai.2023.100124>
- Su, J., & Zhong, Y. (2022). Artificial intelligence (AI) in early childhood education: Curriculum design and future directions. *Computers and Education: Artificial Intelligence*, 3, 100072. <https://doi.org/10.1016/j.caeai.2022.100072>
- Tejedor-García, C. (2020). *Design and evaluation of mobile computer-assisted pronunciation training tools for second language learning* [Tesis doctoral, Universidad de Valladolid]. Universidad de Valladolid. <https://doi.org/10.35376/10324/43663>
- Tejedor-García, C., Escudero-Mancebo, D., Cámara-Arenas, E., González-Ferreras, C., & Cardeñoso-Payo, V. (2020). Assessing pronunciation improvement in students of English using a controlled computer-assisted pronunciation tool. *IEEE Transactions on Learning Technologies*, 13(2), 269–282. <https://doi.org/10.1109/TLT.2020.2980261>
- Uribe-Enciso, O. L., Fuentes-Hernandez, S. S., Vargas-Pita, K. L., & Rey-Pabón, A. S. (2019). Problematic phonemes for Spanish-speakers' learners of English. *GIST – Education and Learning Research Journal*, 19, 215–238. <https://doi.org/10.26817/16925777.701>
- Valente, M. (2023). Peer assessment in the university context for the development of transversal and digital skills. En D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Creative approaches to technology-enhanced learning for the workplace and higher education* (pp. 607–620). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-031-41637-8_49
- Valenzuela, M. G., & French, P. (2023). Production of English vowel contrasts in Spanish L1 learners: A longitudinal study. *Loquens*, 10(1–2), e102. <https://doi.org/10.3989/loquens.2023.e102>
- Zhao, Y. (2024). The exploration of artificial intelligence in pronunciation teaching. En *Proceedings of the 1st International Conference on Data Science and Engineering - Volume 1: ICDSE* (pp. 127–131). SciTePress. <https://doi.org/10.5220/0012829100004547>

Anexos

Anexo 1. Enlaces de modelos creados por los participantes con GTM (URL pública)

<https://teachablemachine.withgoogle.com/models/hCZ2HNAuh/>

<https://teachablemachine.withgoogle.com/models/qik7dGh1I/>

<https://teachablemachine.withgoogle.com/models/DEz2vs-kk/>

<https://teachablemachine.withgoogle.com/models/tFMTVtfFI/>

<https://teachablemachine.withgoogle.com/models/sXAbJGG4p/>

<https://teachablemachine.withgoogle.com/models/Hqr2L38g7/>

<https://teachablemachine.withgoogle.com/models/AXrKKYyfH/>

<https://teachablemachine.withgoogle.com/models/-0P0MAdv6/>

<https://teachablemachine.withgoogle.com/models/GGOVrFaE/>

<https://teachablemachine.withgoogle.com/models/X5SWiThy5/>

<https://teachablemachine.withgoogle.com/models/W7sbALRjp/>

<https://teachablemachine.withgoogle.com/models/S345LKKuP/>

<https://teachablemachine.withgoogle.com/models/LmzUTqWFH/>

<https://teachablemachine.withgoogle.com/models/h2EtS8vGJ/>

<https://teachablemachine.withgoogle.com/models/qLEAx24UT/>

<https://teachablemachine.withgoogle.com/models/6--0r1D0k/>

<https://teachablemachine.withgoogle.com/models/edAu-PQHP/>

<https://teachablemachine.withgoogle.com/models/mDvHrFjEa/>

<https://teachablemachine.withgoogle.com/models/o4Gw33Vnm/>

https://teachablemachine.withgoogle.com/models/1v_6avOOW/

<https://teachablemachine.withgoogle.com/models/KtTyfV1AJ/>

<https://teachablemachine.withgoogle.com/models/EodLnBtii/>

<https://teachablemachine.withgoogle.com/models/MVqFqy5G7/>

Anexo 2. Formato validación expertos para Pretest y Postest

Guía evaluación instrumento

Nombre del evaluador: _____

Último grado de formación: _____

Dirección de correo electrónico: _____

Institución: _____

Cargo que desempeña: _____

Indicaciones: Señor evaluador, se solicita su colaboración para que, después de un riguroso análisis, valide el instrumento dirigido a estudiantes de grado 5to de primaria, con base en su experiencia y criterio profesional en el ámbito de la enseñanza de pronunciación en Inglés en relación con los criterios de claridad, coherencia y pertinencia. Le agradecemos de antemano su amable colaboración.

Anexo a este documento se entregará tanto el Pretest como el Postest propuestos para la investigación.

Para cada criterio considere una escala de 1 a 5, donde:

1: Muy poco

2: Poco

3: Regular

4: Bueno

5: Muy bueno

Objetivo de la investigación: Fortalecer el reconocimiento auditivo y la producción oral de pares mínimos en Inglés en estudiantes de grado quinto del INEM José Celestino Mutis de Armenia mediante una propuesta didáctica basada en la programación de algoriGTMos de clasificación de audio con Google Teachable Machine.

PHONOLOGICAL MINIMAL PAIRS PRETEST

OBJETIVO: Evaluar el reconocimiento auditivo y la producción oral de pares mínimos en Inglés en estudiantes hispanohablantes de nivel A1 – A2 (5TO grado) antes de la intervención.

Área: Inglés

Estructura de la prueba:

- Cada prueba consta de 30 ítems o preguntas divididas en dos secciones o partes.
- Sección A (20 ítems): Reconocimiento Auditivo de Pares Mínimo (Se presenta un formato de selección múltiple con única respuesta).
- Sección B (10 ítems): Producción Oral de Pares Mínimos (Se evalúa mediante la lectura en voz alta o repetición de palabras).

Si el estudiante acierta ambas palabras obtiene 2 puntos, si acierta solamente una palabra obtiene 1 punto y si no acierta ninguna palabra obtiene 0 puntos.

DATOS SOCIODEMOGRÁFICOS:

NOMBRE: _____

EDAD: _____

GÉNERO: Masculino _____ Femenino _____ Otro _____

BARRIO DE RESIDENCIA:

POSEE ACCESO A INTERNET EN CASA: Si _____ No _____

ESTUDIA INGLÉS ATRAVÉS DE ALGÚN CURSO: Si _____ No _____ ¿Cuál?

¿LE GUSTA EL IDIOMA INGLÉS? Si: _____ No _____

DISPOSITIVO CON EL QUE CUENTA EN CASA PARA ESTUDIAR:

Computador _____ Tablet _____ Celular _____

INSTRUCCIONES GENERALES DE LA PRUEBA:

- La prueba está compuesta por dos secciones o partes.
- En la Sección A va a escuchar una serie de palabras en Inglés que el profesor va a ir pronunciando, en su hoja de respuestas debe identificar cuál de las opciones fue la palabra que escuchó y marcar la respuesta correcta. Tenga en cuenta que la palabra será pronunciada en tres ocasiones para que usted rectifique su respuesta antes de marcar ya que no se aceptan tachones o enmendaduras en la prueba.
- En la Sección B debe esperar su llamado por parte del profesor para pronunciar en voz alta algunas palabras en Inglés que se le irán indicando. Tenga en cuenta que debe hacerlo hablando claro y fuerte.

PHONOLOGICAL MINIMAL PAIRS PRETEST**SECCIÓN A: Reconocimiento auditivo**

Ítem	Palabra A	Palabra B
1	Pat	Bat
2	Tip	Dip
3	Coat	Goat
4	Back	Bag
5	Ship	Sheep
6	Bit	Beat
7	Book	Boot
8	Cot	Caught
9	Berry	Very
10	Sip	Zip
11	See	She
12	Ice	Eyes
13	Think	Sink
14	They	Day
15	Thin	Sin

16	Three	Tree
17	Cat	Cut
18	Bed	Bird
19	Sit	Seat
20	Leave	Live

SECCIÓN B: Producción oral

PRETEST

Ítem	Palabra A	Palabra B
1	pat	Bat
2	ship	Sheep
3	berry	Very
4	think	Sink
5	cat	Cut
6	about	Cut
7	tip	Dip
8	book	Boot
9	see	She
10	they	Day

Sección A

CLARIDAD					COHERENCIA					RELEVANCIA				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Observaciones:

Sección B

CLARIDAD					COHERENCIA					RELEVANCIA				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Observaciones:

PHONOLOGICAL MINIMAL PAIRS POSTEST

OBJETIVO: Evaluar el reconocimiento auditivo y la producción oral de pares mínimos en Inglés en estudiantes hispanohablantes de nivel A1 – A2 (5TO grado) después de la intervención.

Área: Inglés

Estructura de la prueba:

- Cada prueba consta de 30 ítems o preguntas divididas en dos secciones o partes.
- Sección A (20 ítems): Reconocimiento Auditivo de Pares Mínimo (Se presenta un formato de selección múltiple con única respuesta).

Sección B (10 ítems): Producción Oral de Pares Mínimos (Se evalúa mediante la lectura en voz alta o repetición de palabras).

Si el estudiante acierta ambas palabras obtiene 2 puntos, si acierta solamente una palabra obtiene 1 punto y si no acierta ninguna palabra obtiene 0 puntos.

NOMBRE: _____

INSTRUCCIONES GENERALES DE LA PRUEBA:

- La prueba está compuesta por dos secciones o partes.
- En la Sección A va a escuchar una serie de palabras en Inglés que el profesor va a ir pronunciando, en su hoja de respuestas debe identificar cuál de las opciones fue la palabra que escuchó y marcar la respuesta correcta. Tenga en cuenta que la palabra será pronunciada en tres ocasiones para que usted rectifique su respuesta antes de marcar ya que no se aceptan tachones o enmendaduras en la prueba.
- En la Sección B debe esperar su llamado por parte del profesor para pronunciar en voz alta algunas palabras en Inglés que se le irán indicando. Tenga en cuenta que debe hacerlo hablando claro y fuerte.

PHONOLOGICAL MINIMAL PAIRS POSTEST**SECCIÓN A: Reconocimiento auditivo**

Ítem	Palabra A	Palabra B
1	pat	Bat
2	tip	Dip
3	coat	Goat
4	back	Bag
5	ship	Sheep
6	bit	Beat
7	book	boot
8	cot	caught
9	berry	very
10	sip	zip
11	see	she
12	ice	eyes
13	think	sink
14	they	day
15	thin	sin
16	three	tree
17	cat	cut
18	bed	bird
19	sit	seat
20	leave	live

SECCIÓN B: Producción oral**POSTEST**

Ítem	Palabra A	Palabra B
1	Pea	Bee
2	Sit	Seat

3	Boat	Vote
4	Thick	Sick
5	Cap	Cup
6	Pull	Pool
7	race	Raise
8	lock	Luck
9	sue	Zoo
10	thigh	Tie

Sección A

CLARIDAD					COHERENCIA					RELEVANCIA				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Observaciones:

Sección B

CLARIDAD					COHERENCIA					RELEVANCIA				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Observaciones:

Comentarios Finales

Dado en _____ a los _____ días del mes de _____ de 2025

Firma del Evaluador _____

Nombre _____

Documento _____

Anexo 3. Formato Consentimiento informado

Consentimiento informado

Introducción

La presente es una solicitud de consentimiento dirigida a padres de familia y/o acudientes de estudiantes de quinto grado de primaria. Este proyecto de investigación busca fortalecer el reconocimiento auditivo y la producción oral de palabras en inglés que suenan muy parecidas (conocidas como *pares mínimos*, es decir, palabras que solo se diferencian por un sonido). El objetivo es ayudar a los estudiantes a mejorar su pronunciación y comprensión auditiva en inglés mediante una propuesta didáctica basada en Google Teachable Machine, una herramienta tecnológica educativa de Google. La participación de su hijo/hija es voluntaria; le invitamos a leer la información a continuación y, si está de acuerdo, a firmar para indicar su consentimiento.

Privacidad, Ética y Consentimiento

- **Participación voluntaria:** La participación en este proyecto es completamente voluntaria y no afectará las evaluaciones ni las calificaciones escolares del estudiante. No habrá ningún tipo de sanción ni consecuencia negativa en caso de no autorizar la participación, y el estudiante puede retirarse del estudio en cualquier momento si así lo desea.
- **Ausencia de riesgos:** La participación no conlleva ningún riesgo para la salud ni el bienestar del estudiante. Todas las actividades se relacionan con el aprendizaje habitual de pronunciación en un entorno controlado y seguro (el aula de clase).
- **Confidencialidad y anonimato:** Todos los datos recopilados (por ejemplo, grabaciones de voz) se manejarán de forma confidencial y anónima. La identidad del estudiante no será divulgada ni aparecerá en los resultados del estudio; las grabaciones y demás datos se identificarán con códigos o seudónimos. Esta información será utilizada únicamente para fines académicos y de investigación, y en ningún caso se compartirá públicamente.
- **Grabación de voz:** Durante el proyecto se realizará la grabación de la voz del estudiante mientras practica las palabras en inglés, previa autorización suya (acudiente). Estas grabaciones de voz se emplearán exclusivamente con fines académicos para analizar los avances en pronunciación y se mantendrán bajo estricta confidencialidad. No se publicarán nombres ni audios que permitan identificar al participante.
- **Gratuidad:** La participación no representa ningún costo para la familia, ni tampoco habrá ningún tipo de remuneración o pago por participar. El beneficio principal es de carácter educativo, ya que el estudiante podrá reforzar sus habilidades de escucha y pronunciación en inglés.

-
- **Almacenamiento de datos:** Los datos recopilados serán almacenados en servidores seguros (Google Drive con acceso restringido). Estos se conservarán únicamente durante el período del estudio y por un máximo de 6 meses tras su finalización. Después de este tiempo, los datos serán eliminados de forma definitiva al igual que la copia de seguridad realizada a estos.
 - **Procedimiento de borrado:** El participante o el acudiente tendrán derecho a solicitar la eliminación de sus datos en cualquier momento. Para hacerlo, deberá comunicarse con el docente investigador a través del correo electrónico alejandranieto199407@gmail.com. En un plazo máximo de 5 días hábiles los registros serán eliminados de forma segura.
 - **Custodia de los audios:** Los archivos de audio estarán bajo custodia del docente titular del grupo y su equipo directo de investigación. Solo estas personas están autorizadas a acceder a ellos. Una vez concluido el estudio y cumplido el periodo de almacenamiento, todos los audios serán eliminados de forma definitiva al igual que la copia de seguridad realizada a estos.
 - **Seguridad adicional y respaldo:** Se han adoptado medidas de protección digital, incluyendo la verificación en dos pasos en la cuenta donde se almacenan los audios, con el fin de garantizar la confidencialidad de la información.
 - **Claridad en los accesos:** Se especifica que únicamente el equipo de investigación directamente involucrado tendrá acceso a los datos. En ningún caso los archivos serán compartidos con terceros sin autorización previa.

Datos sociodemográficos del estudiante

Por favor complete los siguientes datos del estudiante participante:

- **Nombre del estudiante:** _____
- **Edad:** _____ años
- **Género:** _____
- **Barrio de residencia:** _____
- **¿Tiene acceso a internet en casa?** Sí [] No []
- **¿Cuenta con algún dispositivo (computador o celular) en casa?** Sí [] No []

Observaciones (opcional): (En este espacio puede escribir cualquier comentario adicional que considere relevante, ya sea por parte del acudiente o del estudiante.)

Aceptación y firmas

Declaro que he leído y comprendido la información anterior sobre el proyecto, incluyendo su propósito y condiciones éticas. En consecuencia, doy mi consentimiento para que mi hijo/hija participe en esta investigación y autorizo la grabación de su voz para los fines académicos señalados.

- **Nombre del acudiente (padres de familia o tutor):**

- **Firma del acudiente:** _____
- **Firma del estudiante:** _____
- **Fecha:** _____

Anexo 4. Formato prueba Pretest

PHONOLOGICAL MINIMAL PAIRS PRETEST

OBJETIVO: Evaluar el reconocimiento auditivo y la producción oral de pares mínimos en Inglés en estudiantes hispanohablantes de nivel A1 – A2 (5TO grado) antes de la intervención.

Área: Inglés

Estructura de la prueba:

- Cada prueba consta de 30 ítems o preguntas divididas en dos secciones o partes.
- Sección A (20 ítems): Reconocimiento Auditivo de Pares Mínimo (Se presenta un formato de selección múltiple con única respuesta).
- Sección B (10 ítems): Producción Oral de Pares Mínimos (Se evalúa mediante la lectura en voz alta o repetición de palabras).

DATOS SOCIODEMOGRÁFICOS:

NOMBRE: _____

EDAD: _____

GÉNERO: Masculino _____ Femenino _____ Otro _____

BARRIO DE RESIDENCIA: _____

POSEE ACCESO A INTERNET EN CASA: Si _____ No _____

ESTUDIA INGLÉS ATRAVÉS DE ALGÚN CURSO: Si _____ No _____ ¿Cuál?

¿LE GUSTA EL IDIOMA INGLÉS? Si: _____ No _____

DISPOSITIVO CON EL QUE CUENTA EN CASA PARA ESTUDIAR:

Computador _____ Tablet _____ Celular _____

INSTRUCCIONES GENERALES DE LA PRUEBA:

- La prueba está compuesta por dos secciones o partes.
- En la Sección A va a escuchar una serie de palabras en Inglés que el profesor va a ir reproduciendo, en su hoja de respuestas debe identificar cuál de las opciones fue la palabra que escuchó y marcar la respuesta correcta. Tenga en cuenta que la palabra será pronunciada en tres ocasiones para que usted verifique su respuesta antes de marcar ya que no se aceptan tachones o enmendaduras en la prueba.

- En la Sección B debe esperar su llamado por parte del profesor para pronunciar en voz alta algunas palabras en Inglés que se le irán indicando. Tenga en cuenta que debe hacerlo hablando claro y fuerte.

PHONOLOGICAL MINIMAL PAIRS PRETEST

SECCIÓN A: Reconocimiento auditivo

Ítem	Palabra A	Palabra B
1	Pat	bat
2	Tip	dip
3	Coat	goat
4	Back	bag
5	Ship	sheep
6	Bit	beat
7	Book	boot
8	Cot	caught
9	Berry	very
10	Sip	zip
11	See	she
12	Ice	eyes
13	Think	sink
14	They	day
15	Thin	sin
16	Three	tree
17	Cat	cut
18	Bed	bird
19	Sit	seat
20	Leave	live

SECCIÓN B: Producción oral**PRETEST**

Ítem	Palabra A	Palabra B
1	pat	bat
2	ship	sheep
3	berry	very
4	think	sink
5	cat	cut
6	cap	cup
7	tip	dip
8	book	boot
9	see	she
10	they	day

Anexo 5. Formato prueba Postest

PHONOLOGICAL MINIMAL PAIRS POSTEST

OBJETIVO: Evaluar el reconocimiento auditivo y la producción oral de pares mínimos en Inglés en estudiantes hispanohablantes de nivel A1 – A2 (5TO grado) después de la intervención.

Área: Inglés

Estructura de la prueba:

- Cada prueba consta de 30 ítems o preguntas divididas en dos secciones o partes.
- Sección A (20 ítems): Reconocimiento Auditivo de Pares Mínimo (Se presenta un formato de selección múltiple con única respuesta).
- Sección B (10 ítems): Producción Oral de Pares Mínimos (Se evalúa mediante la lectura en voz alta o repetición de palabras).

NOMBRE: _____

INSTRUCCIONES GENERALES DE LA PRUEBA:

- La prueba está compuesta por dos secciones o partes.
- En la Sección A va a escuchar una serie de palabras en Inglés que el profesor va a ir pronunciando, en su hoja de respuestas debe identificar cuál de las opciones fue la palabra que escuchó y marcar la respuesta correcta. Tenga en cuenta que la palabra será pronunciada en tres ocasiones para que usted rectifique su respuesta antes de marcar ya que no se aceptan tachones o enmendaduras en la prueba.
- En la Sección B debe esperar su llamado por parte del profesor para pronunciar en voz alta algunas palabras en Inglés que se le irán indicando. Tenga en cuenta que debe hacerlo hablando claro y fuerte.

PHONOLOGICAL MINIMAL PAIRS POSTEST**SECCIÓN A: Reconocimiento auditivo**

Ítem	Palabra A	Palabra B
1	Pat	bat
2	Tip	dip
3	Coat	goat
4	Back	bag
5	Ship	sheep
6	Bit	beat
7	Book	boot
8	Cot	caught
9	Berry	very
10	Sip	zip
11	See	she
12	Ice	eyes
13	Think	sink
14	They	day
15	Thin	sin
16	Three	tree
17	Cat	cut
18	Bed	bird
19	Sit	seat
20	Leave	live


SECCIÓN B: Producción oral**POSTEST**

Ítem	Palabra A	Palabra B
1	Pat	bat
2	Ship	sheep
3	Berry	very
4	Think	sink
5	Cat	cut
6	Cap	cup
7	Tip	dip
8	Book	boot
9	See	she
10	They	day

Anexo 6. Formato encuesta de motivación estudiante / docente

Encuesta de motivación. Nombre: _____

- Marca con una X el número que mejor describa cómo te sientes
- 1 = Nada, 2 = Poco, 3 = Regular, 4 = Mucho, 5 = Muchísimo

Espacio solo para el docente 

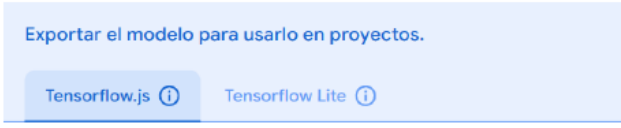
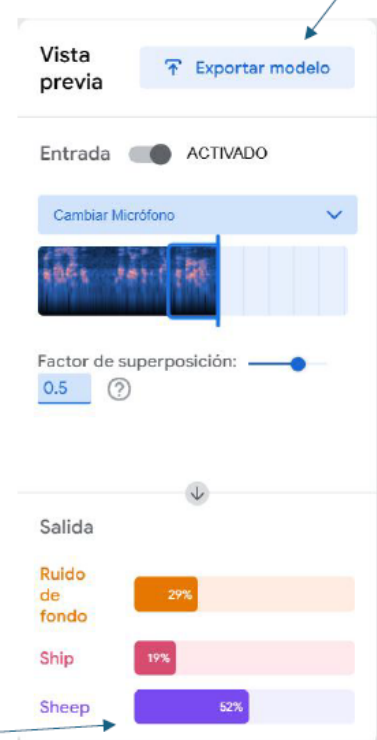
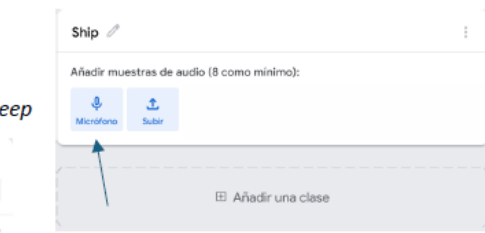
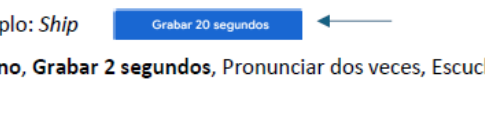
Ítem	Pregunta	1	2	3	4	5
1	Me divertí durante la actividad.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Pude realizar la actividad.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Quiero seguir practicando.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	El modelo me ayudó a escuchar mejor.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Me gustó trabajar en equipo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Comentarios:					

#	Descripción	S / N
1	Participación alta	
2	Persistencia	
3	Colaboración	
4	Enfoque	
5	Frustración notable	
6	Iniciativa	

Anexo 7. Formato guía rápida de Google Teachable Machine (GTM)

Guía Rápida de Google Teachable Machine (GTM)

1. Abrir el navegador (Google) y buscar: **Google Teachable Machine**
2. Clic al primer enlace: **teachablemachine.withgoogle.com**
3. Clic al botón de **Primeros pasos**
4. Clic a **Proyecto de audio**
5. Grabar el Ruido de Fondo: Clic en **Micrófono**, Hacer silencio, **Grabar 20 segundos**, Esperar, Clic en **Extraer muestra**
6. Cambiar nombre a las clases: Clic en **Class 2**, Escribir, Por ejemplo: **Ship**
7. Grabar nuevamente, pronunciando la palabra: Clic en **Micrófono**, **Grabar 2 segundos**, Pronunciar dos veces, Escuchar, Clic en **Extraer muestra**.
8. Repetir el paso 7 al menos cuatro veces. Si después de escuchar, cometen errores, no dan clic en Extraer muestras sino en **Grabar 2 segundos**. También pueden eliminar audios si dan clic sobre el **bote de basura** al poner el mouse sobre ellos.
9. Clic en **Añadir una clase**.
10. Repetir pasos 6, 7, 8. Pero con la otra palabra, Por ejemplo: **Sheep**
11. Repetir hasta añadir todos los Pares mínimos.
12. Clic en **Preparar modelo**.
13. Probar el modelo.
14. Anotar los datos obtenidos. Pronuncia la palabra y anota el número que dice el modelo. Por ejemplo: **Ship – 97% – 89%**
15. Después de escribir las palabras pronunciadas y los valores de precisión. Exporta el modelo.
16. Clic en **Subir (enlace para compartir)**.
17. Clic en **Subir mi modelo**.
18. Copiar el enlace para compartir.
19. Pega el enlace en un documento para no perderlo. Guárdalo.



Esta es la parte importante

Anexo 9. Formato de encuesta postintervención GE

ENCUESTA PARA ESTUDIANTES

Proyecto de pronunciación en inglés con Teachable Machine

Instrucciones

Marca con una X tu respuesta. Recuerda que no hay respuestas buenas o malas.

1. ¿Te gustaron las actividades que hicimos en clase? ¿Por qué?

2. ¿Sentiste que las actividades te ayudaron a mejorar tu pronunciación en inglés?

Sí

No

3. ¿Crees que ahora puedes pronunciar mejor algunas palabras en inglés que antes?

Sí

No

4. ¿La herramienta que usamos (Teachable Machine) fue fácil de utilizar con ayuda del profe?

Sí

No

5. ¿Volverías a hacer actividades como estas en tus clases de inglés? ¿Por qué?

6. ¿Qué tanto disfrutaste la intervención?

(Encierra la carita que mejor te represente)

😊 Mucho

🙂 Un poco

😐 Más o menos

😞 Poco

😓 Nada

7. Si pudieras cambiar algo de las actividades, ¿qué cambiarías?

Anexo 10. Formato de encuesta postintervención GC

Encuesta para estudiantes – GC

Clases de pronunciación en inglés

Instrucciones

Marca con una X tu respuesta. Recuerda que no hay respuestas buenas o malas.

1. ¿Te gustaron las actividades de pronunciación que hicimos en clase? ¿por qué?

2. ¿Sentiste que estas actividades te ayudaron a mejorar tu pronunciación en inglés?

Sí

No

3. ¿Crees que ahora puedes pronunciar mejor algunas palabras en inglés que antes?

Sí

No

4. ¿Las actividades de pronunciación fueron fáciles de entender y seguir con ayuda del profe?

Sí

No

5. ¿Volverías a hacer este tipo de actividades en tus clases de inglés?

Sí

No

6. ¿Qué tanto disfrutaste la actividad?

7. Si pudieras cambiar algo de las actividades, ¿qué cambiarías?

Anexo 11. Bitácora de IA

El desarrollo de la presente investigación integró de manera responsable diversas herramientas de IA como apoyo instrumental en las etapas de redacción, búsqueda bibliográfica, análisis y verificación, sin sustituir la interpretación crítica ni la originalidad de los investigadores. En primer lugar, el equipo de investigación realizaba sesiones de lluvia de ideas para delimitar los puntos centrales de cada sección del documento. A partir de estos insumos preliminares, se formulaban preguntas específicas que orientaban la búsqueda de información en bases de datos académicas (ScienceDirect, Google Académico) y mediante el uso de plataformas de apoyo como SciSpace, que facilitó la exploración de artículos científicos y la interacción con los documentos en formato PDF.

Con las fuentes seleccionadas, se recurrió a NotebookLM y SciSpace para localizar pasajes relevantes y sintetizar respuestas a preguntas concretas, mientras que los borradores producidos se enriquecieron con la asistencia de ChatGPT, herramienta utilizada para organizar los contenidos de forma lógica, mejorar la redacción académica, eliminar redundancias y sugerir observaciones críticas. Posteriormente, cada propuesta generada era verificada de manera manual mediante la confrontación con los documentos originales y la revisión en grupo.

Asimismo, ChatGPT se utilizó en la codificación en Python para realizar cálculos estadísticos y generar representaciones gráficas, lo cual permitió complementar los análisis cuantitativos. En paralelo, el chat generativo de SciSpace y las búsquedas asistidas en NotebookLM sirvieron para identificar estudios adicionales que respaldaran los ajustes propuestos en el documento.

Durante todo el proceso, se aplicaron verificaciones adicionales. En particular, se empleó el detector de IA de SciSpace, cuya pertinencia radicó en su orientación académica y en la membresía institucional disponible. Sin embargo, se constató que la herramienta presentaba limitaciones, al señalar como “generado por IA” textos elaborados directamente por los autores o al emitir reportes imprecisos en etapas posteriores. Esta experiencia evidenció que no resulta responsable delegar a la IA la “humanización” de un texto, pues esta tarea corresponde a la revisión crítica de los investigadores.

Finalmente, se complementó la revisión formal con el uso de Turnitin para garantizar la detección de similitudes y la prevención del plagio. De este modo, se configuró un ecosistema de

uso de IA que apoyó la eficiencia de la redacción y el análisis, sin comprometer la originalidad del trabajo.

En la Tabla A1 se documenta con detalle ejemplos de la fecha, herramienta, prompt, salida generada, uso en la tesis y procedimiento de verificación aplicado en cada caso, constituyéndose en un instrumento de transparencia y trazabilidad.

Tabla 6 Bitácora de uso de IA

Fecha	Herramienta	Prompt	Salida	Uso	Verificación
11/06/2025	SciSpace	“AI-based tools like Google Teachable Machine enhance phonological learning and pronunciation of minimal pairs in primary education”	Búsqueda profunda de artículos en diferentes bases de datos con información de sus resultados, métodos, contribuciones, etc.	Utilizado principalmente en el planteamiento del problema y el marco teórico.	Seguimiento de las referencias, búsqueda de los artículos en sus fuentes originales, descarga de PDF para su análisis con NotebookLM
03/08/2025	SciSpace	“Phonological contrasts in English challenging for Spanish speakers”	Generación de un reporte referenciado sobre dificultades fonológicas	Fundamental para el diseño metodológico y el marco teórico.	Seguimiento de las referencias, búsqueda de los artículos en sus fuentes originales, descarga de PDF para su

					análisis con NotebookLM
13/08/2025	NotebookLM	“Encuentra en las fuentes anexas información relacionada con...”	Extracto referenciado de los PDFs	Implementado para el diseño de la metodología y el marco teórico.	Verificado con lectura directa del PDF, siguiendo las referencias generadas
20/08/2025	ChatGPT	“Esta es una lista de preguntas y respuestas relacionadas con la fonética, toma la siguiente estructura y distribuye la información en un borrador de la X del marco teórico”	Texto generado con IA, un borrador para su posterior refuerzo	Este proceso se siguió para casi la totalidad de las secciones. Se le daba a ChatGPT unos huesos y él ayudaba a crear el esqueleto. Primordial para el marco teórico, metodología y resultados.	Corrección de estilo manual, verificación con el verificador de IA de SciSpace (Aunque como se menciona en la declaración de uso de IA, esto no fue práctico)
14/09/2025	NotebookLM	“Toma el siguiente fragmento y referéncialo con	Extracto referenciado de los PDFs	Fundamental para la redacción del marco teórico	Corrección de estilo manual

		las fuentes anexas. Qué ideas relacionadas se exploran en adición”			
20/09/2025	ChatGPT	“Crea un código Python Google Colab para organizar en dataframes los resultados pre y post con cálculo de las siguientes métricas.”	Código en Python con análisis estadístico y gráfica	Fundamental para la presentación de los resultados	Cotejado con manual en Excel y verificación de las funciones con sus respectivas librerías
22/09/2025	ChatGPT	“Redacta observaciones críticas para mejorar discusión, resultados, etc.”	Párrafo sugerido	Como insumo para nutrir el texto ya redactado	Contrastado y descartado por redundante
22/09/2025	Turnitin	Reporte de similitud	Informe de coincidencias	Sí, validación general del manuscrito	Revisión de coincidencias por autores