Diseño e implementación de un robot interactivo como herramienta tecnológica de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años, mediante actividades interactivas de comprensión de cuentos, formación de oraciones y corrección de pronunciación.

Cabezas Guevara, Jefferson Javier y Moreno Guijarro, Kevin Mateo

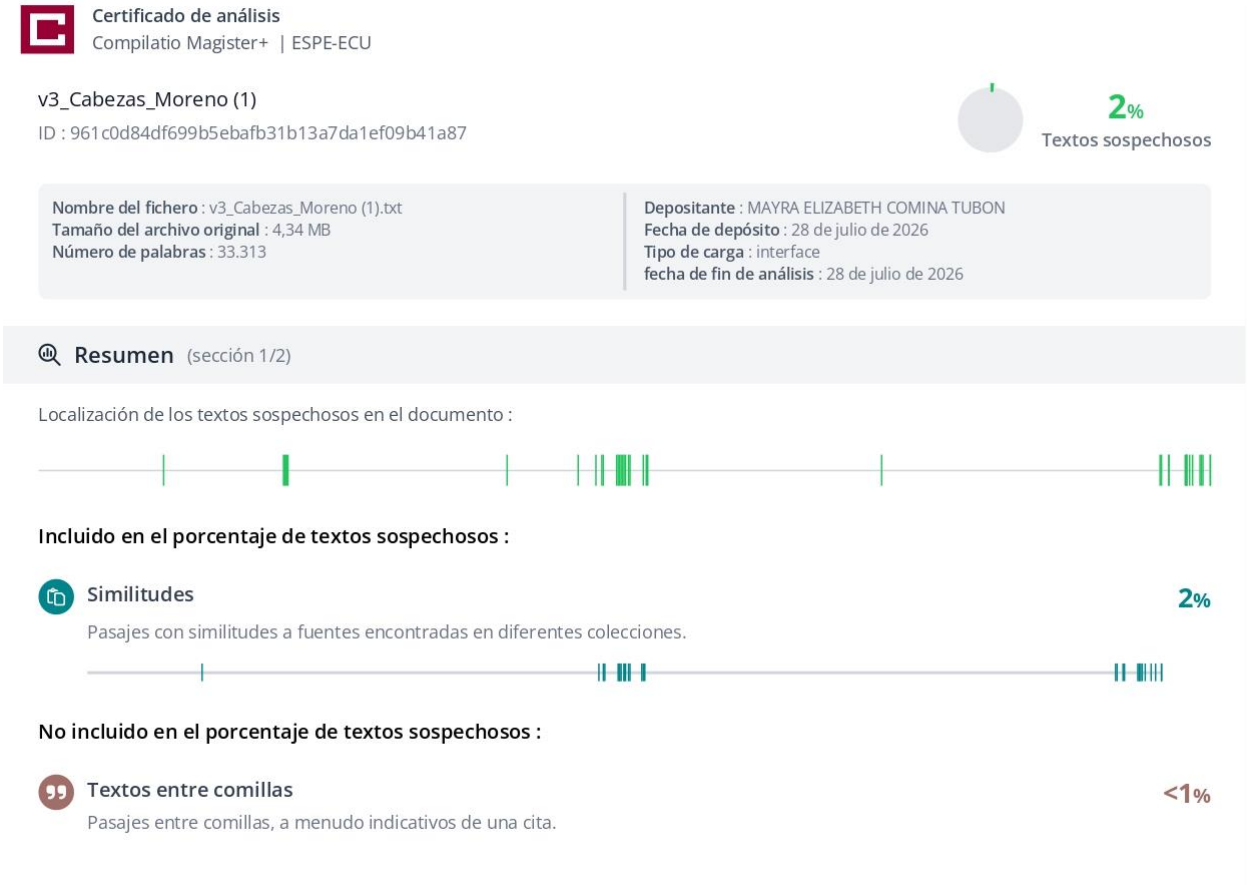
Departamento de Ciencias de la Energía y Mecánica

Carrera de Mecatrónica

Trabajo de integración curricular, previo a la obtención del título de Ingeniero en Mecatrónica

Ing. Comina Tubón Mayra Elizabeth.Ph.D.

27 de julio de 2026



Certificado de análisis
Compilatio Magister+ | ESPE-ECU

v3_Cabezas_Moreno (1)
ID : 961c0d84df699b5ebafb31b13a7da1ef09b41a87

2%
Textos sospechosos

Nombre del fichero : v3_Cabezas_Moreno (1).txt
Tamaño del archivo original : 4,34 MB
Número de palabras : 33.313

Depositante : MAYRA ELIZABETH COMINA TUBON
Fecha de depósito : 28 de julio de 2026
Tipo de carga : interface
fecha de fin de análisis : 28 de julio de 2026

Resumen (sección 1/2)

Localización de los textos sospechosos en el documento :

Incluido en el porcentaje de textos sospechosos :

Similitudes 2%
Pasajes con similitudes a fuentes encontradas en diferentes colecciones.

No incluido en el porcentaje de textos sospechosos :

Textos entre comillas <1%
Pasajes entre comillas, a menudo indicativos de una cita.

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
MAYRA ELIZABETH
COMINA TUBON

Ing. Comina Tubón Mayra Elizabeth, Ph.D.

Directora

Departamento de Ciencias de la Energía y Mecánica

Carrera de Mecatrónica

Certificación

Certifico que el trabajo de integración curricular: "Diseño e implementación de un robot interactivo como herramienta tecnológica de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años, mediante actividades interactivas de comprensión de cuentos, formación de oraciones y corrección de pronunciación." fue realizado por los señores Cabezas Guevara, Jefferson Javier y Moreno Guijarro, Kevin Mateo, el mismo que cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE; habiéndose efectuado una revisión integral de forma y de fondo conforme a los lineamientos y formatos institucionales vigentes. Además, fue revisado y analizada en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

Sangolquí, 27 de julio del 2026

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
MAYRA ELIZABETH
COMINA TUBON

Ing. Comina Tubón Mayra Elizabeth, Ph.D.

Directora

C.C.: 1724994684

Departamento de Ciencias de la Energía y Mecánica

Carrera de Mecatrónica

Responsabilidad de Autoría

Nosotros, Cabezas Guevara, Jefferson Javier y Moreno Guijarro, Kevin Mateo, con cédulas de ciudadanía n.° 2100753181 y 1750169573, declaramos que el contenido, ideas y criterios del trabajo de integración curricular: “Diseño e implementación de un robot interactivo como herramienta tecnológica de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años, mediante actividades interactivas de comprensión de cuentos, formación de oraciones y corrección de pronunciación.” es de nuestra autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 27 de julio del 2026

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
JEFFERSON JAVIER CABEZAS GUEVARA

Cabezas Guevara, Jefferson Javier

C.C.: 2100753181

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
KEVIN MATEO MORENO GUIJARRO

Moreno Guijarro, Kevin Mateo

C.C.: 1750169573

Departamento de Ciencias de la Energía y Mecánica

Carrera de Mecatrónica

Autorización de Publicación

Nosotros, Cabezas Guevara, Jefferson Javier y Moreno Guijarro, Kevin Mateo, con cédulas de ciudadanía n.° 2100753181 y 1750169573, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de integración curricular: "Diseño e implementación de un robot interactivo como herramienta tecnológica de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años, mediante actividades interactivas de comprensión de cuentos, formación de oraciones y corrección de pronunciación." en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

Sangolquí, 27 de julio del 2026

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
JEFFERSON JAVIER CABEZAS GUEVARA

Cabezas Guevara, Jefferson Javier

C.C.: 2100753181

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
KEVIN MATEO MORENO GUIJARRO

Moreno Guijarro, Kevin Mateo

C.C.: 1750169573

Dedicatoria

A mi madre Gloria, por levantarme cuando pensé que no podría más, a mi padre Calistrato, por darme las bases para defenderme en cualquier ámbito, a mi hermano Jerson, por guiarme en las decisiones que para mí resultaban difíciles, a mi hermana Yalid, por ser mi mejor amiga y enseñarme como ser yo mismo, a mi mascota Kira por ser el apoyo emocional que necesitaba en mi adolescencia.

Además, a los amigos que hice en la universidad, que en mis peores momentos no dudaron en ser mi soporte para poder llegar juntos a esta meta.

CABEZAS GUEVARA, JEFFERSON JAVIER

Dedico este trabajo a las personas que son el motor de mi vida: mi familia. A mis padres, Julio y María, gracias por enseñarme que el sacrificio siempre vale la pena y que el mayor tesoro es la familia. Papi, gracias por cada esfuerzo para que nunca nos faltara nada. Mami, gracias por cada palabra de amor y fortaleza. A mi hermano Diego, gracias por ser mi compañero y ejemplo. A mi novia Katty, gracias por caminar conmigo. A mi perrito Loki. Todo lo que soy lleva una parte de ustedes. Con todo mi corazón, este logro también les pertenece.

MORENO GUIJARRO, KEVIN MATEO

Agradecimiento

Agradezco a Dios por acompañarme y bendecirme en cada paso de este camino. A toda mi familia, especialmente a mis abuelitos Carmen, Lucho, Judy y Pepe, por su amor y apoyo incondicional. A quienes hoy me cuidan desde el cielo: mi abuelito Lucho, mi tío Chuchito, mi tía Liliam, mi amigo Alan y mi perrito Fito; su recuerdo sigue dándome fuerzas para seguir adelante. Gracias a mis amigos del colegio y de la universidad por hacer más llevaderos los momentos difíciles. A todos quienes formaron parte de este proceso, gracias por creer en mí y acompañarme hasta este logro.

MORENO GUIJARRO, KEVIN MATEO

Índice de contenido

Certificación 3

Responsabilidad de Autoría 4

Autorización de Publicación 5

Dedicatoria 6

Agradecimiento 7

Índice de contenido 8

Índice de Figuras 12

Índice de Tablas 15

Resumen 17

Abstract 18

Capítulo I 19

Antecedentes 19

Justificación e importancia 21

Alcance del proyecto 23

Objetivos 29

Objetivo General 29

Objetivos Específicos 30

Capítulo II 31

Estado del Arte 31

Educación infantil y tecnologías interactivas 31

Importancia del desarrollo lingüístico en la primera infancia 31

Tendencias en la integración de tecnologías emergentes en educación infantil 34

Robots narradores y aprendizaje basado en cuentos.....37

Narración asistida por robots: fundamentos teóricos y pedagógicos 37

Efectos de la emoción y la expresividad robótica en la motivación y aprendizaje 43

Comparativa entre narración digital tradicional y narración con robots 45

Interacción humano-robot para el aprendizaje lingüístico.....48

Modelos de interacción niño-robot en contextos educativos 49

Personalización y adaptación emocional del robot 51

Tecnologías de reconocimiento y retroalimentación lingüística 53

Reconocimiento automático del habla infantil (ASR) 53

Síntesis de voz en robots educativos 55

Retroalimentación de pronunciación y corrección fonética automatizada 57

Integración multimodal: voz, visión y gestos para la interacción natural 59

Revisión de estudios y análisis de tendencias 61

Revisión sistemática de la robótica social en educación infantil 61

Análisis bibliométrico de la investigación sobre robots sociales.....63

Brechas y oportunidades para el diseño de robots lingüísticos.....65

Capítulo III 73

Diseño Conceptual.....73

Matriz QFD 74

Generación de Conceptos 83
Concepto 1 83
Concepto 2 84
Diseño del Prototipo 87
Sistema Mecánico 87
Sistema Electrónico y de Control 99
Diseño PCB 111
Interfaz Gráfica 115
Desarrollo de Software 128
Fabricación y ensamblaje del prototipo 131
Reconocimiento de voz 135
Captura de audio 141
Generación de voz 143
Administrador de Contenido 145
Capítulo IV 151
Pruebas y resultados 151
Pruebas del Sistema Mecánico 151
Pruebas de precisión angular en: Expresión de felicidad 151
Pruebas de precisión angular en: Expresión de tristeza 153
Pruebas de la interfaz gráfica 155
Validación del prototipo final 157

Análisis de Costos 157

Validación mediante encuestas a expertos 161

Capítulo V 165

Conclusiones y recomendaciones 165

Conclusiones generales 165

Recomendaciones técnicas y operativas 167

Trabajos Futuros 168

Bibliografía 169

Apéndices 176

Índice de Figuras

Figura 1 *Esquema de Tecnologías*.....25

Figura 2 *Resumen gráfico de la propuesta.*.....28

Figura 3 *Esquema de Navegación de la Interfaz.*.....28

Figura 4 *Robot NAO.*.....33

Figura 5 *Matriz QFD*.....75

Figura 6 *Sistema de movimiento de 2 grados de libertad*.....79

Figura 7 *Movimientos anatómicos de la cabeza.*.....81

Figura 8 *Referencia del concepto 1.*.....84

Figura 9 *Referencia del concepto 2.*.....85

Figura 10 *Sistema de movimiento de brazo.* 88

Figura 11 *Sistema de movimiento de cabeza.*.....91

Figura 12 *Tensión máxima que soporta el eje de la cabeza.*.....93

Figura 13 *Deformación máxima ejercida en el eje de la cabeza.*.....93

Figura 14 *Coeficiente de Seguridad mínimo en el eje de la cabeza.* 94

Figura 15 *Torso que soporta los esfuerzos principales* 95

Figura 16 *Tensión máxima soportada por el Torso* 96

Figura 17 *Deformación máxima en el Torso*.....96

Figura 18 *Factor de Seguridad mínimo en el Torso* 97

Figura 19 *Cabeza con visualización de sus matrices led* 98

Figura 20 *Visualización general del robot*.....99

Figura 21 *Diagrama General de Conexiones* 101

Figura 22 *Actuadores utilizados* 104

Figura 23 *Consumo de voltaje y corriente*.....105

Figura 24 *Conexiones del sistema Lógico*.....108

Figura 25 *Conexiones del Sistema de Potencia* 109

Figura 26 *Diseño PCB para el proyecto.* 112

Figura 27 *Diseño de pistas PCB.* 113

Figura 28 *Conexiones para diseño de PCB.* 114

Figura 29 *Diagrama General de la navegación de la HMI.* 115

Figura 30 *Pantalla de menú de juegos de la HMI* 116

Figura 31 *Pantalla de menú de cuentos de la HMI* 117

Figura 32 *Ventana de reproducción de cuentos de la HMI* 118

Figura 33 *Pantalla de preguntas para cuentos* 119

Figura 34 *Pantalla de felicitaciones de respuestas correctas.* 120

Figura 35 *Pantalla de menú de selección de oraciones.* 121

Figura 36 *Pantalla de Ordenar Oraciones.* 122

Figura 37 *Pantalla de juego de ordenar las oraciones.* 123

Figura 38 *Pantalla del juego de Pronunciación.* 124

Figura 39 *Pantalla de saludo del módulo Platica con el robot.* 125

Figura 40 *Expresiones aproximadas del Robot.* 127

Figura 41 *Prueba de funcionamiento mecánico preliminar.* 129

Figura 42 *Prueba de funcionamiento mecánico preliminar.* 130

Figura 43 *Posición de impresión del eje principal del cuello.* 131

Figura 44 *Configuración del STL de la cabeza.* 132

Figura 45 *Prueba de funcionamiento mecánico preliminar.* 133

Figura 46 *Post-procesado del robot.* 134

Figura 47 *Diseño final del robot.* 135

Figura 48 *Resumen del proceso de reconocimiento de voz.* 138

Figura 49 *Proceso de pronunciación.* 140

Figura 50 *Proceso de pronunciación incorrecta.* 142

Figura 51 *Resumen del Reconocimiento de Voz*.....145

Figura 52 *Pantalla de inicio del Administrador de Contenido.*.....146

Figura 53 *Pantalla de administración de Cuentos.* 147

Figura 54 *Formulario de edición de un cuento.* 148

Figura 55 *Pantalla de administración de Oraciones.* 149

Figura 56 *Pantalla de administración de categorías de Pronunciación.*.....150

Figura 57 *Pruebas en la expresión de felicidad*.....152

Figura 58 *Resumen de resultados en la secuencia de expresión de felicidad*.....153

Figura 59 *Pruebas en la expresión de tristeza* 154

Figura 60 *Resumen de resultados en la secuencia de expresión de tristeza.* 155

Figura 61 *Tiempos de carga de las pantallas del HMI*.....156

Figura 62 *Resultados de reconocimiento de voz*.....157

Figura 63 *Resumen general de costos*.....159

Figura 64 *Comparativa de precios con prototipos comerciales* 161

Figura 65 *Distribución de respuestas — Educación Inicial*.....162

Figura 66 *Distribución de respuestas — Ingeniería (Energía y Mecánica)* 163

Figura 67 *Distribución de respuestas — Total General (ambas encuestas)* 164

Índice de Tablas

Tabla 1 *Resumen de Trabajos consultados* 67

Tabla 2 *Necesidades del robot*.....73

Tabla 3 *Necesidades del robot*.....76

Tabla 4 *Módulo 1 (Material General)* 77

Tabla 5 *Material para impresión 3D* 77

Tabla 6 *Módulo 2 (Tarjeta de control principal)* 77

Tabla 7 *Tarjeta de control secundaria* 78

Tabla 8 *Módulo 3 (Visualización)* 78

Tabla 9 *Modelo de pantalla* 78

Tabla 10 *Módulo 5 (Movimiento de brazo): Tecnología* 78

Tabla 11 *Módulo 5 (Movimiento de brazo): Grados de libertad* 79

Tabla 12 *Motor para el movimiento de brazo* 80

Tabla 13 *Módulo 6 (Movimiento de cabeza): Tecnología* 80

Tabla 14 *Módulo 6 (Movimiento de cabeza): Grados de libertad*.....81

Tabla 15 *Motor para el movimiento de cabeza*.....81

Tabla 16 *Módulo 7 (Sistema de Ojos)* 82

Tabla 17 *Panel digital para los ojos*.....82

Tabla 18 *Módulo 7 (Fuente de alimentación)* 82

Tabla 19 *Valorización de conceptos*.....86

Tabla 20 *Selección de conceptos*.....86

Tabla 21 *Arquitectura de comunicación.* 127

Tabla 22 *Resumen de resultados de movimiento en la expresión de felicidad.* 152

Tabla 23 *Resumen de resultados de movimiento en la expresión de tristeza.* 154

Tabla 24 *Costos de Hardware.*.....158

Tabla 25 *Costos Mecánicos*.....158

Tabla 26 *Costos Electrónicos de Desarrollo*.....158

Tabla 27 *Costos de Fabricación de PCB*.....159

Resumen

El presente trabajo de integración curricular aborda el diseño e implementación de un robot interactivo educativo orientado a apoyar el desarrollo de habilidades lingüísticas en niños de 4 a 5 años mediante actividades de comprensión de cuentos, formación de oraciones y corrección de pronunciación. Se abordó el proyecto de acuerdo con la metodología de diseño mecatrónico VDI 2206 a partir de la revisión del estado del arte de más de 35 referencias de campo sobre robótica social, narración asistida y reconocimiento del habla, en el que se pudo constatar la inexistencia de soluciones actualmente existentes para este grupo de edad. Posteriormente, la matriz QFD priorizó la visualización gráfica, la interacción mediante el tacto y el movimiento de los brazos y/o de la cabeza como los tres conceptos en liza; de dos conceptos fue el Concepto 1 el que gano (52/100 frente a 43) debido a su coste y durabilidad. Se construyó un prototipo que utilizó manufactura aditiva FDM en PLA y donde los esfuerzos máximos alcanzaron 0.366 MPa y 0.138 MPa en la cabeza y el torso, respectivamente. El control electrónico está formado por una Raspberry Pi 5 y una ESP32, utilizando una fuente de alimentación de 5V y 15A y teniendo unos factores de seguridad de 5.51 y 1.03. El software cuenta con una interfaz gráfica compuesta por 14 pantallas, reconocimiento de voz (Google/Vosk) y síntesis de voz con Piper. Las pruebas concluyeron con un 100 % de fiabilidad mecánica, un tiempo de carga de la interfaz de 0,86 s y una pronunciación de un 95,7 % sobre 1.000 evaluaciones. El costo total del prototipo fue de 630,20 USD. De las encuestas de validación cualitativa llevadas a cabo con cuatro docentes de Educación Inicial y cinco docentes de Ingeniería se encuentran promedios de 4,75 y 4,57 (sobre 5) y un 70,2 % de respuestas de Totalmente de acuerdo, que validan la viabilidad técnica, económica y pedagógica del sistema como herramienta de apoyo para el desarrollo del lenguaje.

Palabras clave: Robótica educativa, robot social, reconocimiento de voz, narración interactiva, inteligencia artificial educativa, interfaz multimodal.

Abstract

The work discussed in this curriculum integration project is about the design and implementation of an interactive robot that has been created to help the child aged between 4 to 5 years develop language skills through activities including story comprehension, sentence building, and pronunciation correction. The project was developed according to the VDI 2206 mechatronic design framework, based on a review of the state of the art of over 35 references on social robotics, assisted storytelling, and speech recognition. This review revealed a lack of solutions currently designed for this age group. Subsequently, a QFD matrix was used to prioritize graphical visualization, touch-based interaction, and arm and/or head movement. Of the two concepts, Concept 1 was selected (52/100 vs. 43) as a result of its cost and life expectancy. Using FDM printing with PLA, a prototype has been made, ensuring highest values of stress equal to 0.366 MPa and 0.138 MPa for the head and body, respectively. The electronic unit of the system is made up of Raspberry Pi 5 and ESP32 powered by a power source of 5V and 15A, with safety factors of 5.51 and 1.03. Operational software has a GUI and features above-mentioned 14 screens and Google/Vosk recognition and Piper’s sound synthesis. Testing resulted in 100% mechanical efficiency, 0.86 second interface loading time, and 95.7% of words correctly pronounced out of 1,000 measurements. Overall cost of the prototype is 630.20 USD. Validation surveys conducted with four pre-school teachers and five engineering university professors resulted in averaging 4.75 and 4.57 scores of 5. 70.2% of the entire responses stated "Strongly agree." These results validate the system's technical, economic, and pedagogical viability as a support tool for language development in early childhood.

Keywords: Educational robotics, social robot, voice recognition, interactive storytelling, educational artificial intelligence, multimodal interface.

Capítulo I

En este capítulo se abordará la fundamentación teórica y contextual del proyecto, explorando los principales antecedentes que sustentan el desarrollo de un robot educativo interactivo orientado al fortalecimiento de las habilidades lingüísticas en la primaria. Con la intención de acotar la situación presente y dar una base a la realización de este trabajo se realizará una revisión de la bibliografía propia de la robótica educativa, la narración asistida por software y el reconocimiento de voz. Con ello, se revisarán detalladamente los logros conseguidos en relación con estas tecnologías, así como los problemas y las oportunidades que presentan. Dicho análisis se llevará a cabo sin perder de vista la importancia que la IA tiene en la interacción socioemocional en las configuraciones educativas.

Antecedentes

La robótica educativa, experimental hace apenas una década, funciona hoy como herramienta pedagógica para reforzar el aprendizaje del lenguaje desde la primera infancia (Chatzichristofis, 2023). Los robots sociales —diseñados con un comportamiento y una apariencia similares a la de un compañero de juego— promueven la interacción verbal de los niños, tal y como atestigua la evidencia existente. Kory y Breazeal (2014) lo pusieron a prueba de manera directa al dejar que un robot contase cuentos y viendo cómo los niños incorporaban léxico y patrones narrativos nuevos únicamente por jugar con él. El seguimiento a largo plazo ratificó la tendencia, pues después de varias sesiones de juego e interacción los niños empezaban a imitar el léxico y el estilo lingüístico del robot, con mejoras objetivas en su propio lenguaje (Kory-Westlund & Breazeal, 2019). Essien y Parbanath (2024) legaron a la misma conclusión cuando sometieron a revisión la literatura reciente sobre el ámbito de los cuentos mediadas por un robot: este tipo de mediación produce mejor atención de los alumnos y produce también un mejor rendimiento de los aprendizajes en comparación con una narración

convencional. En definitiva, estos resultados sugieren que el intercambio verbal entre el niño y el robot, siempre que se produzca en un contexto lúdico, favorece el desarrollo del lenguaje.

Las tecnologías que soportan esta aplicación han ido avanzando con el mismo paso. Los estudios sobre robots sociales en la primera infancia abordan pocas veces la alfabetización temprana de forma directa, pero sí respaldan el papel de los robots en el aprendizaje del lenguaje oral en preescolar (Neumann, 2020). El reconocimiento de voz, por su parte, ganó precisión en los últimos años para transcribir el habla infantil. Janssens et al. (2024) señalan que modelos de aprendizaje profundo como Whisper transcriben correctamente cerca del 60% de los enunciados producidos por niños (Janssens et al., 2024), se ha llegado a un nivel suficientemente alto para soportar una interacción en tiempo real. Asimismo, se han probado algoritmos de retroalimentación de pronunciación: Spaulding et al. (2018) elaboraron un sistema en que un robot evaluaba los fonemas de los niños gracias a la API SpeechAce y canalizaba el contenido del juego de palabras de acuerdo con los resultados que obtenía. Todo esto proporciona base para saber que un robot educativo, que narra cuentos, genera preguntas guiadas y es capaz de detectar errores de pronunciación.

Planteamiento del problema

La robótica educativa tiene como elemento común en el aprendizaje infantil el juego, la interacción y la tecnología; en los últimos años ha ido aterrizando en las aulas. En el caso de la educación infantil básica, la comunicación y pronunciación, su aplicación es todavía escasa. La mayoría de la investigación se centra en la robótica social y en el reconocimiento de la voz para la educación primaria o para el aprendizaje de lenguas extranjeras, dejando de lado la aplicación sistemática y continuada de robots como mediadores lingüísticos en la educación infantil preescolar. Esa falta de experiencias diseñadas para niños de 4 a 5 años es, al mismo tiempo, una oportunidad y un reto pedagógico y tecnológico.

A esa edad los niños aprenden rápido, y el contexto de comunicación junto con la retroalimentación inmediata resultan determinantes para consolidar habilidades lingüísticas sólidas. Los materiales de enseñanza tradicionales no siempre logran captar el interés de los pequeños en la expresión oral o la pronunciación.

Los nuevos avances en la identificación automática del lenguaje infantil, así como en los modelos de aprendizaje profundo, como son los casos de Whisper o de SpeechAce, suponen nuevas oportunidades aparecidas para que una máquina dé voz y respuesta al lenguaje verbal de los niños en el mismo momento. No obstante, la evidencia acerca de ese uso específico —en el que se produce en un contexto de historia a la carta, donde el robot no sólo habla, sino que también escucha, reacciona y promueve la práctica del lenguaje verbal por el niño— sigue siendo escasa. De ahí la cuestión central de la presente investigación: ¿de qué maneras podría un robot educativo interactivo mejorar las habilidades lingüísticas de los niños de 4 a 5 años?

La carencia de medios robóticos que fomenten el desarrollo del lenguaje en la infancia de forma lúdica y con validación experimental radica el problema. Para cerrar la brecha, hay que cruzar la ingeniería mecatrónica, la inteligencia artificial y la pedagogía del lenguaje a través de un solo proyecto de investigación aplicada, en este caso aplicado, que sea capaz de medir qué tan bien puede llegar a ser un robot en la educación infantil.

Justificación e importancia

La presente tesis procura dar respuesta a la necesidad de disponer de tecnologías interactivas que favorezcan el aprendizaje significativo del lenguaje desde la más tierna infancia. El resultado de diferentes investigaciones es coincidente en que los robots sociales incrementan los niveles de motivación, sostienen la atención, y favorecen la predisposición a la comunicación verbal de los niños, elementos fundamentales para el desarrollo lingüístico infantil. En comparación con los recursos audiovisuales, inertes y neutros, un robot proporciona interacción

multimodal (verbal, de gestos, de mirada o responder físicamente), que se aproxima en lejanía a la propia dinámica del lenguaje del ser humano, lo que facilita el aprendizaje de vocabulario nuevo, la imitación del habla y la corrección de la pronunciación en una situación atractiva para el niño.

El interés del proyecto se ubica, sobre todo, en su perfil interdisciplinario, en el que conversan robótica educativa, terapia del habla y didáctica del lenguaje. La finalidad de este proyecto viene condicionada por la elaboración de un prototipo de robot educativo que actúe como narrador interactivo, explorando historias, haciendo preguntas y ofreciendo retroalimentación verbal y visual con carácter de tiempo real, no solo enseñando vocabulario y estructuras lingüísticas sino también favoreciendo el desarrollo socio-emocional del niño que desarrolle la empatía, pueda expresar emociones y desarrolle la confianza en la comunicación verbal. Además, apunta al mismo objetivo que las tendencias internacionales actuales de innovación tecnológica en educación que promueven el uso de robots-compañeros de aprender en diseños inclusivos.

Desde un punto de vista técnico, el presente trabajo articula un prototipo robótico innovador, funcional y de bajo coste; sus algoritmos de reconocimiento y síntesis de voz junto a sensores visuales y táctiles integrados a un modelo mecatrónico intentan dar respuesta a las características cognitivas y emocionales de los niños pequeños, teniendo visión de futuras aplicaciones escolares y terapéuticas. La propuesta aporta de este modo a la investigación científica y a la práctica educativa a la vez que ofrece evidencia empírica del efecto de la robótica en el desarrollo del lenguaje.

La relevancia social del estudio se relaciona con un objetivo que tiene que ver con la democratización del acceso a tecnologías educativas emergentes. En gran parte de Latinoamérica las herramientas de robótica aun no se encuentran adaptadas ni en el idioma ni

en lo que significa pedagógicamente para sus aulas; este proyecto pretende cubrir esa brecha y mostrar que se puede construir herramientas tecnológicas adecuadas pedagógicamente y pertinentes culturalmente. La tesis deja, de este modo, conocimiento científico y un recurso práctico para fortalecer las competencias comunicativas de los niños desde sus primeros años de vida.

Alcance del proyecto

Esta sección detalla el alcance del proyecto: las etapas, los recursos y las tecnologías que permitirán llevar a cabo el diseño, desarrollo e implementación del robot educativo interactivo. Se definen las fases metodológicas, desde la investigación preliminar y el diseño mecánico-electrónico hasta la validación experimental del prototipo, junto con las herramientas de hardware y software que sustentan su funcionamiento. Se delimitan también las condiciones, limitaciones y resultados esperados del estudio, y con ellos el marco operativo con el que se evaluará la efectividad del sistema propuesto para fortalecer las habilidades lingüísticas en niños de 4 a 5 años.

Etapas del Proyecto

El desarrollo del proyecto se estructurará en las siguientes fases:

1. Investigación preliminar: revisión bibliográfica sobre robots sociales en educación infantil y sobre las tecnologías actuales de reconocimiento de voz y narración interactiva; incluye identificar los requerimientos específicos para la interacción lingüística con niños de 4 a 5 años.

2. Diseño del sistema: diseño de los planos que disponen la estructura mecánica del robot, selección de los componentes, actuadores, display y sensores, y selección de las bibliotecas de software para voz, imagen y diálogo.

3. Desarrollo de la interfaz de usuario: creación de una interfaz visual y acústica por la que el niño efectúe la interacción con el sistema, incluida la configuración correspondiente de los periféricos de entrada y salida.

4. Fabricación e integración del prototipo: ensamblar el robot su base, sus uniones y su electrónica, ajustar el sistema para el reconocimiento de la voz, para emitir los audios de los cuentos y las preguntas para la retroalimentación.

5. Pruebas y validación: evaluar el sistema en los escenarios de test simulados, ejercicios de detección de voz, de comprensión del preguntador, de corrección de errores de pronunciación y del resultado de estos cambios.

6. Documentación y conclusiones: confección de un reporte técnico del diseño, desarrollo e instalación del prototipo, con el seguimiento de los resultados experimentales y las conclusiones sobre su impacto educativo.

Tecnologías Utilizadas

El proyecto utilizará tecnologías como se muestra en la Figura 1, propias de la ingeniería mecatrónica y la computación interactiva aplicadas al desarrollo del robot lingüístico. Comprenden:

1. Hardware: micrófono y cámara como entrada principal de información, servomotores para brazos y cabeza, estructura impresa en 3D, pantalla táctil de 10 pulgadas y módulo de sonido con altavoces.

2. Software: softwares de reconocimiento automático de voz (ASR) con adaptación a voz infantil, síntesis de voz (TTS), procesamiento básico de imagen con seguimiento facial, controladores de los servomotores, interfaz gráfica infantil (GUI) y motor de diálogo interactivo basado en reglas.

3. Manufactura: impresión 3D para las partes estructurales del robot, corte láser para detalles estéticos o funcionales y mecanizado simple de la parte estructural del robot para ensamblar articulaciones.

Figura 1
Esquema de Tecnologías

Nota. Elaboración propia.

Resultados Esperados

Los principales entregables del proyecto serán:

- Un prototipo funcional de robot interactivo, capaz de narrar cuentos, formular preguntas comprensivas y evaluar la pronunciación infantil.
- Se elabora un informe técnico que expone a nivel detallado el diseño, el desarrollo, la integración y la validación del sistema.
- Datos experimentales en relación al rendimiento del robot ya sea en cuanto a las métricas de reconocimiento de voz, en cuanto a la detección de errores de pronunciación, o en cuanto a la receptividad de los niños durante la interacción con el robot.

Limitaciones

La pieza en la que se centra el trabajo es para la población infantil que va de los 4 a 5 años se desarrolla en español. El reconocimiento de voz ve su uso limitado únicamente a la identificación de palabras o de frases cortas relacionadas con los cuentos que se encuentran predefinidos, sin la capacidad de realizar ningún tipo de aprendizaje automático profundo o personalización de la lengua a un nivel más individualizado. El robot representa un prototipo a escala pequeña de uso experimental (no está sistematizado como para llevarse a la industrialización y comercialización inmediata), aunque sí que podrá ser ampliado a otros rangos etarios, lenguas u otras capacidades de inteligencia artificial avanzada.

Recursos Necesarios

El desarrollo del proyecto precisará acceso a:

- Laboratorios: Espacios habilitados para poner a prueba estaciones de trabajo de programación, herramientas de ensamblaje electrónico, equipos de impresión 3D y corte láser para la confección de prototipos.
- Software especializado: programas de software CAD para diseño mecánico, entornos para aplicaciones de desarrollo integradas (IDE) para la programación del microcontrolador, bibliotecas para el reconocimiento de voz y la síntesis, herramientas para el diseño de las interfaces de usuario.
- Materiales: motores de corriente continua, pantallas digitales, Raspberry PI o microcontroladores similares, micrófonos de alta sensibilidad, cámaras web, altavoces, bancos mecánicos de la estructura y baterías extra.

Partes Interesadas (Stakeholders)

Los principales agentes implicados en el proyecto son:

- Institución académica: el departamento de ingeniería y los docentes tutores que siguen el proceso de la construcción del prototipo y la investigación pedagógica vinculada a la propuesta de proyecto.
- Instituciones educativas preescolares: Los jardines infantiles y los centros educativos que podrían hacer uso del robot como herramienta de ayuda en actividades relacionadas con el lenguaje.
- Familias y educadores: Los padres y los educadores que quieren incorporar tecnologías innovadoras que favorezcan el desarrollo temprano del lenguaje en los niños.

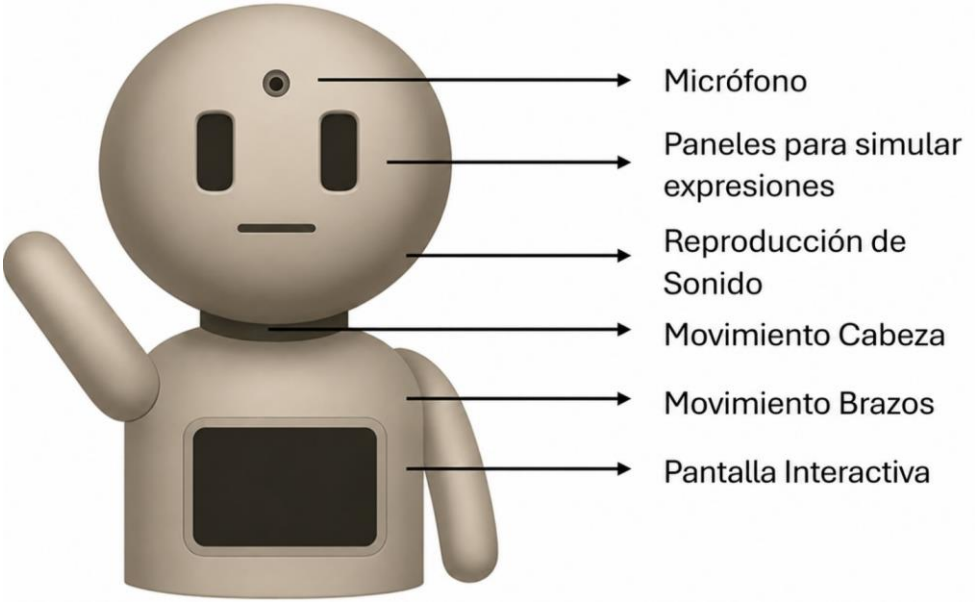
Áreas de trabajo

Dentro del proyecto se trabaja en tres principales ejes, que son:

- Mecánica: Diseño estructural del robot con piezas móviles en brazos y cabeza, soporte para pantalla táctil e integración física de sensores y actuadores. Fabricación realizada con impresión 3D y montaje manual.
- Electrónica: Configuración del sistema de control (Raspberry Pi), servomecanismos, detectores (micrófono) y efectores. Integrandose de manera sincronizada.
- Control y programación: Comprende el desarrollo de software para comunicar los diferentes subsistemas del robot como són: control de movimiento, reconocimiento de voz, síntesis verbal, lógica de preguntas y retroalimentación, control de pantalla, y cuentos interactivos.

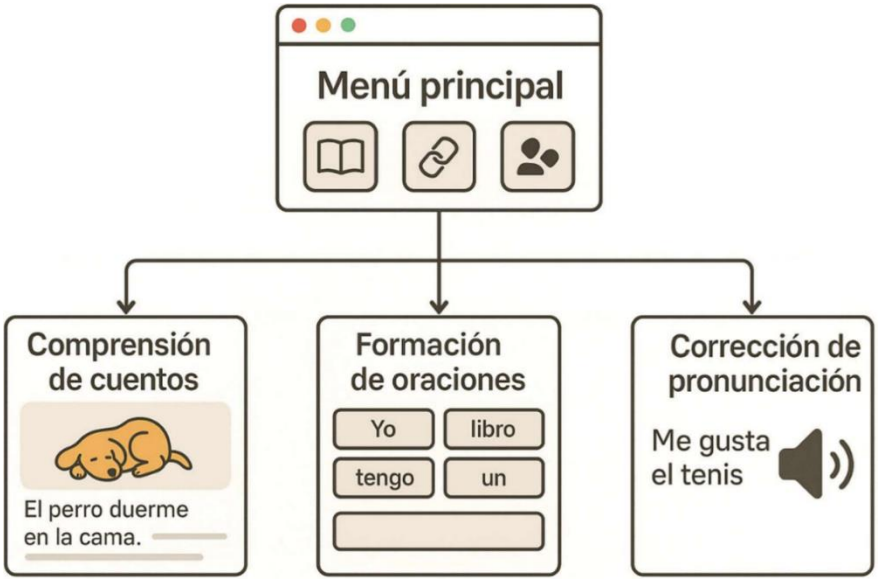
En la Figura 2 y Figura 3 se puede tener un resumen de la propuesta para la realización del robot y la navegación que tendrá para el aprendizaje de los niños.

Figura 2
Resumen gráfico de la propuesta.



Nota. Elaboración propia.

Figura 3
Esquema de Navegación de la Interfaz.



Nota. Elaboración propia.

Limitaciones

El ámbito se considera restringido al trabajo con niños de 4 a 5 años y a la lengua española, dado que la participación se ha limitado a un reconocimiento de voz restringido a palabras y frases simples y a los cuentos fijados en la aplicación. El sistema no hace uso de aprendizaje automático profundo ni de la personalización del lenguaje individualizada. El robot se mantiene en la fase de prototipo a escala reducida, de uso experimental, sin ninguna intención de entrar en producción masiva ni comercialización inmediata. La extensión a otros rangos etarios, idiomas o capacidades de inteligencia artificial avanzada queda como trabajo futuro.

Objetivos

En esta sección se presentan el objetivo general y los objetivos específicos que guían el desarrollo del proyecto, los cuales orientan de manera estructurada el diseño e implementación de un robot educativo interactivo destinado a niños de 4 a 5 años. La determinación de dichos objetivos sienta las bases para el desarrollo de una revisión del estado del arte y del diseño mecánico, electrónico y de control del prototipo, a la par del desarrollo de la interfaz gráfica y la posterior validación experimental del sistema. Por tanto, el conjunto de estos objetivos va a marcar el alcance y la dirección metodológica seguir en el proyecto, ya que cada una de las fases, va a servir para ayudar a la mejora de las habilidades lingüísticas de la población infantil a la que se dirige el proyecto mediante el uso de tecnología robótica.

Objetivo General

Diseñar e implementar un robot interactivo como herramienta tecnológica de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años, mediante actividades interactivas de comprensión de cuentos, formación de oraciones y corrección de pronunciación.

Objetivos Específicos

- Realizar un estudio del arte para el diseño y construcción de un robot interactivo de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años
- Diseñar el sistema mecánico, electrónico y de control del robot interactivo de apoyo para el desarrollo de habilidades lingüísticas en niños de 4 a 5 años
- Desarrollar la interfaz gráfica en la que se muestren juegos interactivos que ayudarán a la enseñanza de habilidades lingüísticas de niños de 4 a 5 años.
- Implementar los sistemas diseñados para el funcionamiento del prototipo.
- Realizar la validación y pruebas de funcionamiento del prototipo desarrollado.

Capítulo II

Estado del Arte

Este capítulo presenta la revisión bibliográfica sobre la robótica educativa y su relación con el desarrollo de habilidades lingüísticas en la infancia. Se analizan los avances en robótica social, narración asistida por robots y tecnologías de reconocimiento de voz aplicadas al aprendizaje del lenguaje, con el fin de identificar tendencias, aportes y vacíos que fundamenten el diseño del robot interactivo propuesto en esta investigación.

Educación infantil y tecnologías interactivas

Esta parte trata sobre la relación que existe entre el mundo de la educación infantil y el uso de las tecnologías interactivas herramientas como la robótica educativa, la inteligencia artificial y la realidad aumentada, que innovan los procesos relacionados con la enseñanza y el aprendizaje en las etapas tempranas de la infancia. Se estudia el papel que juega la importancia del desarrollo lingüístico en edades tempranas y el rol de los robots sociales como mediadores del aprendizaje, que ayudan a fomentar la comunicación, la motivación y la interacción emocional con los niños. Además también se abordan las tendencias actuales de la incorporación de tecnologías emergentes en el entorno educativo, poniendo el enfoque en su potencial para la creación de experiencias multifacéticas, personalizadas y afectivas que contribuyan a un desarrollo cognitivo y lingüístico integral.

Importancia del desarrollo lingüístico en la primera infancia

La robótica educativa es un área interdisciplinaria que une principios de ingeniería, informática, pedagogía y psicología para fomentar habilidades cognitivas, motrices y sociales mediante la interacción con robots. Esta área ha evolucionado de simples kits de programación y ensamblaje hacia sistemas inteligentes los cuales son capaces de comunicarse, percibir información y adaptarse al usuario. Según Chatzichristofis (2023), los avances recientes en

robótica educativa han permitido que los robots se conviertan, de herramientas mecánicas de exploración tecnológica a mediadores sociales y cognitivos dentro del aula, promoviendo experiencias de aprendizaje dinámicas y personalizadas.

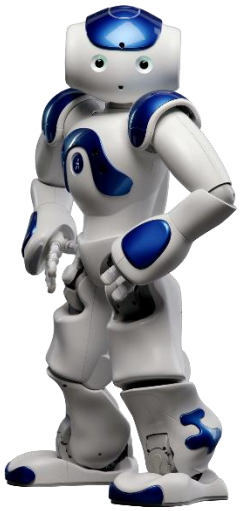
Al inicio del desarrollo de la robótica educativa, esta tenía el desarrollo de habilidades lógico-matemáticas y de resolución de problemas centradas en diseñar y programar robots simples. Pero la llegada de la robótica social modificó esta línea de trabajo hacia robots que son capaces de comunicarse con los alumnos mediante el uso de lenguaje, gestos y expresiones emocionales. En este sentido, Breazeal et al. (2016) explican que este cambio de paradigma es resultado de que se considera que los robots pueden ser usados como una herramienta técnica, pero también como una herramienta social, lo cual influye sobre la motivación, la atención y la colaboración de los alumnos. La robótica educativa moderna pretende emular aspectos de la comunicación humana para generar entornos de aprendizaje más inmersivos y afectivos.

El Springer Handbook of Robotics expone el hecho de que la robótica social está considerada una importante evolución dentro de la robótica educativa, ya que incorpora la interacción humano-robot (HRI) y la cooperación adaptativa a modo de ejes de diseño. Un robot social, desde este punto de vista, no solamente ejecuta tareas, sino que también ha de interpretar emociones, responder a señales no verbales y ser capaz de mantener de forma coherente una interacción a lo largo del tiempo. Si lo trasladamos al escenario del aula ello quiere decir que el robot pasa a ser un compañero de aprendizaje que ajusta el comportamiento a las necesidades del estudiante, al tiempo que se favorece el aprendizaje colaborativo y el refuerzo positivo (Breazeal et al., 2016).

Tal y como lo menciona Fink (2012) defiende la idea de que la medida de semejanza con el ser humano - lo que denomina human likeness - va a afectar el grado de aceptación, confianza y efectividad pedagógica que tendrá el robot. Las investigaciones demuestran que los niños se

implican más cuando perciben rasgos humanos en el robot: voz expresiva, movimientos naturales o expresiones faciales básicas. Este antropomorfismo positivo ha orientado el diseño de numerosos robots educativos, desde prototipos sencillos hasta modelos avanzados como el robot NAO, que se muestra en la Figura 4, junto con otros como Pepper o Kaspar Figura 4(Fink, 2012).

Figura 4
Robot NAO.



Nota. Obtenido de (Nomura, 2017).

Hoy la robótica educativa se entiende como una herramienta integral que combina aprendizaje activo, interacción social y pensamiento computacional. En el trabajo presentado por Chatzichristofis (2023) se presenta una investigación más reciente en el mundo en el que se incorporan sensores multimodales, la inteligencia artificial, y el aprendizaje automático para que el robot puede leer el contexto en el que se encuentra el entorno educativo y cambiar las respuestas en el tiempo real; es decir, el robot podrá leer la atención del niño, detectar emociones básicas, dar pie a saber cuándo modular la voz del robot según la actividad y, por lo tanto ello le permitirá ampliar su uso a otros campos como el desarrollo del lenguaje, la alfabetización inicial y la educación inclusiva.

La robótica educativa dejó de dirigirse a "aprender a programar" para orientar su foco hacia "aprender con robots": lo que prima hoy en día es la interacción, la comunicación y la adaptación del aprendizaje. Breazeal et al. (2016) reclaman que el objetivo no es "enseñar" contenidos académicos sino generar vivencias que susciten curiosidad, empatía y colaboración. Este cambio viene dado por la confluencia de tres factores: un hardware que se ha vuelto cada vez más accesible, una evolución continua en inteligencia artificial y, por último, un acercamiento a una mejor comprensión del rol del afecto en el aprendizaje.

Este recorrido muestra un tránsito de propuestas puramente tecnológicas hacia modelos centrados en lo social y lo emocional. Fink (2012), Breazeal et al. (2016) y Chatzichristofis (2023) coinciden en que el futuro de la disciplina apunta a robots pedagógicos empáticos, capaces de adaptarse a distintos estilos de aprendizaje, dar retroalimentación inmediata y participar activamente en el desarrollo cognitivo y lingüístico infantil. Esta base conceptual sustenta el diseño del robot interactivo propuesto en la presente investigación, pensado para integrar esas dimensiones humanas y tecnológicas y favorecer así el aprendizaje del lenguaje en edades tempranas.

Tendencias en la integración de tecnologías emergentes en educación infantil

En las últimas dos décadas la escena educativa ha cambiado con rapidez debido a la incorporación de nuevas tecnologías. Entre ellas, la realidad aumentada (RA), la inteligencia artificial (IA) y la robótica educativa destacan por su impacto en la manera de enseñar y en la forma en que participan los estudiantes. En la etapa infantil, estas herramientas abren la posibilidad de crear experiencias interactivas y adaptativas que apoyan el desarrollo cognitivo, lingüístico y socioemocional desde edades tempranas (Pasalidou et al., 2025).

Pasalidou et al. (2025) identifican como una de las tendencias más relevantes la combinación de realidad aumentada con robótica educativa para impulsar la exploración activa

y un aprendizaje con sentido para el niño. Esta integración permite manipular objetos virtuales sobre el entorno físico al mismo tiempo que se interactúa con un robot real, de forma colaborativa, lo que refuerza habilidades motoras y espaciales y tiende un puente entre el pensamiento abstracto y lo concreto. Con estas experiencias llamadas "phygital" (física y digital a la vez) produce el cambio de un aprendizaje simbólico a un aprendizaje práctico, haciendo visible y tangible la naturaleza de las consecuencias de las acciones que realizan, tanto en el mundo físico como en el mundo virtual.

La realidad aumentada ha mostrado especialmente el modo en que puede mantener la atención de los niños por mucho tiempo en actividades educativas, ya que sus recursos visuales e interactivos les permite representar ideas abstractas mediante imágenes o simulaciones; algo que es muy útil en preescolar, etapa donde los niños desarrollan el pensamiento concreto y aún no han adquirido el simbólico. Pasalidou et al. (2025) subrayan además que al combinar RA con robots sociales el aprendizaje adquiere una dimensión emocional más marcada, ya que el niño empieza a ver al robot no solo como un dispositivo, sino como un compañero que lo acompaña a explorar mundos virtuales. Esa combinación estimula la curiosidad, la creatividad y la memoria visual, elementos que resultan importantes en las primeras etapas de adquisición del lenguaje.

La inteligencia artificial opera, en cambio, de forma transversal: personaliza el contenido y las interacciones según lo que hace cada estudiante. En la enseñanza a nivel infantil, los sistemas de IA detectan pautas de comportamiento, valoran el grado de comprensión e introducen una retroalimentación orientada. Pasalidou et al. (2025) indican que los algoritmos aplicados a robots o plataformas de RA pueden modificar la dificultad de una actividad, atenuar el tono de voz del robot o seleccionar narraciones que se correspondan con el nivel lingüístico del niño/a, lo que genera una experiencia de aprendizaje más inclusiva y respetuosa respecto de cada joven aprendiz.

La tendencia general, por tanto, apunta hacia la complementariedad de estas tecnologías bajo un enfoque de aprendizaje multisensorial y multimodal. En los proyectos más avanzados —que combinan visión por computadora, reconocimiento de voz y aprendizaje automático— los robots ya son capaces de reconocer expresiones faciales y objetos a través de cámaras para acompañar a los niños en su desarrollo del habla. Esta convergencia tecnológica ha dado lugar a espacios educativos inteligentes en los que robots y RA responden de manera coherente a las emociones y comportamientos infantiles, reforzando así su motivación y su sentido de autonomía.

La literatura nueva también manifiesta un creciente interés por los sistemas educativos híbridos donde el robot hace la media entre el niño y los contenidos digitales. En estos contextos la RA envuelve personajes o entornos que forman parte de la interacción con el robot físico generando nuevas formas de narración aumentada, en las que el niño realiza este proceso de creación de narraciones, realiza las preguntas o ejecuta los movimientos que el robot sugiere, se produce la confluencia entre la exploración tecnológica y la imaginación del niño. Esta aplicación es especialmente interesante para la enseñanza de lenguas en cuanto que promueve la comprensión auditiva, la expresión oral y la relación de palabras con las representaciones u acciones de estas.

De acuerdo con Pasalidou y colaboradores (2025), el potencial inherente a la RA, la IA y la robótica educativa reside en la posibilidad de crear entornos adaptativos, cooperativos y multisensoriales en línea con los principios del constructivismo pedagógico. No obstante, alerta que ese potencial únicamente se da en la medida en que se realice una planificación pedagógica y se lleve a cabo una formación del profesorado tal que se garantice el uso de dichas herramientas de forma efectiva y ética.

La educación infantil deja atrás, en conjunto, los modelos de enseñanza pasivos y avanza hacia paradigmas de aprendizaje interactivo y personalizado, impulsados por estos desarrollos tecnológicos, en los que la pedagogía funciona como puente entre el niño y la tecnología. La combinación de RA, IA y robótica educativa redefine el rol del niño, que pasa a ser un agente activo y co-creador de conocimiento: un cambio de paradigma real en la forma de entender la educación infantil en la era digital.

Robots narradores y aprendizaje basado en cuentos

En esta sección del estudio se examina el papel de los robots narradores y el aprendizaje mediante cuentos como herramientas educativas en la enseñanza infantil. Se examina cómo la narración con robots, al integrar lenguaje, emoción, interacción social y aprendizaje, potencia el desarrollo del lenguaje y la comprensión narrativa en niños, además de presentar los fundamentos teóricos y pedagógicos que respaldan el uso de robots como herramientas de enseñanza, considerando también su influencia en motivación, atención y empatía durante las narraciones. Al final se compara la narración digital tradicional con la narración robótica, y se destacan las ventajas de esta última para crear experiencias educativas más dinámicas, afectivas y personalizadas, en las que el niño participa activamente en la construcción de significado.

Narración asistida por robots: fundamentos teóricos y pedagógicos

Contar cuentos es uno de los métodos más antiguos, y a la vez más eficaces, del aprendizaje infantil, porque conjuga lo emocional con la comprensión del lenguaje y la construcción de significado. En su versión más reciente, esta práctica se apoya en la narración digital y, más recientemente, en la narración guiada por robots, donde un robot social funciona como narrador interactivo o compañero de aprendizaje. Este enfoque se apoya en las teorías del aprendizaje socioconstructivista y en prácticas pedagógicas narrativas que priorizan el diálogo,

la imitación y la construcción conjunta del conocimiento a partir de experiencias compartidas (Kory & Breazeal, 2014).

El estudio pionero de Kory y Breazeal (2014) mostró que los robots narradores tienen potencial como recurso de aprendizaje de una lengua, especialmente para preescolares. En dicho estudio, un robot leía historias interactivas a niños de esta edad, que tomaban turnos para comunicarse mientras que el robot daba vida a la actividad realizando preguntas y utilizando gestos. Los resultados reflejan una mejoría significativa en el vocabulario de los niños y en el dominio de la estructura narrativa; la evidencia empírica que respalda la idea de que el aprendizaje de la lengua mejora cuando el niño concibe al robot como un compañero comunicador y no meramente como un recurso didáctico.

Partiendo de la teoría, esta modalidad de contar historias está bien sustentada por los principios del aprendizaje social de Bandura y por el concepto de apego comunicativo, ya que el niño aprende mediante la imitación del comportamiento lingüístico del adulto con el que él o ella tiene algún tipo de relación. Al interactuar el robot con una respuesta verbal coherente, unas expresiones faciales adecuadas y unos gestos ajustados al contexto, se establece una dinámica de conversación análoga a la humana con el fin de modular el habla y comprender las reglas sintácticas. Esto no es un proceso mecánico: ocurre dentro de un marco de juego y afecto que refuerza la motivación intrínseca del niño para comunicarse.

El efecto de la emoción y la congruencia afectiva entre el robot narrador y el contenido de la historia fue el foco de un estudio posterior de Appel et al. (2021). Observaron que los niños se sumergen más en la narración y desarrollan una actitud más positiva hacia el robot cuando la expresión emocional acompaña la trama, por ejemplo con una voz triste en los pasajes tristes o un tono alegre en los momentos felices. Llamaron a esta condición congruencia afectiva: sugiere que la eficacia educativa de un robot narrador no depende solo de la fluidez del habla, sino

también del tono de voz, la expresión facial y el movimiento corporal en sintonía con la emoción de la historia. La emoción, de hecho, actúa como mediadora cognitiva que favorece tanto la retención de información como la empatía hacia lo narrado.

El mismo trabajo de Appel et al. (2021) aporta además un sustento pedagógico relevante desde la teoría del transporte narrativo, según la cual las historias emocionalmente coherentes facilitan el procesamiento cognitivo y el aprendizaje implícito. Aplicada a la robótica educativa, esta idea sugiere que el compromiso emocional con la historia condiciona no solo lo que el niño percibe, sino cómo responde a ella. En ese sentido, el robot deja de ser un simple transmisor de información y pasa a mediar experiencias comunicativas que integran emoción y cognición, un aspecto central en el desarrollo del lenguaje infantil.

Essien y Parbanath (2024) ofrecen, por su parte, una visión más amplia sobre las tendencias y modelos educativos de la narración asistida por robots: señalan que esta práctica combina elementos de la narración tradicional, la educación mediada por tecnología y la interacción social. En su revisión proponen una clasificación de los enfoques existentes divididos en tres categorías:

1. Robots narradores autónomos: son robots que cuentan cuentos de forma autónoma y hacen preguntas sencillas.
2. Sistemas colaborativos: en el que el robot y el niño co-construyen la historia de una manera turnada o mediante decisiones conjuntas.
3. Entornos híbridos: la narración es complementada por recursos digitales como imágenes, proyecciones o realidad aumentada.

Según los autores Essien y Parbanath (2024), de los tres modelos detectados, el enfoque colaborativo es el que mayores beneficios ha mostrado en cuanto a la educación infantil de la

primera infancia, dado que favorece tanto la participación como la autorregulación emocional, que son dos de las habilidades que están relacionadas con la alfabetización temprana. Los robots narradores, señalan, mejoran la comprensión lectora y la memoria verbal al construir un ambiente lúdico y de bajo riesgo donde el error se asume como parte natural del aprendizaje.

La lógica de este tipo de narración se sustenta tanto en la teoría del aprendizaje experiencial como en el constructivismo social, y ayuda al niño a explorar el lenguaje mediante la interacción misma. El robot funciona, en cierto modo, como un andamio digital que transmite comprensión a través de retroalimentación inmediata y apoyo emocional, algo particularmente valioso en la primera infancia, cuando los niños están aprendiendo a organizar oraciones, distinguir emociones y seguir el hilo de una historia. Un robot expresivo, además, ayuda a transmitir la habilidad de comunicarse en un entorno humano, lo que refuerza la capacidad lingüística del niño.

Kory y Breazeal (2014), Appel et al. (2021) y Essien y Parbanath (2024) coinciden, en definitiva, en situar la narración asistida por robots dentro de un enfoque pedagógico que integra lenguaje, emoción y tecnología: en la educación infantil temprana, esta no solo refuerza el desarrollo del lenguaje, sino que también estimula la empatía, la escucha y la comprensión simbólica. Estas tres investigaciones respaldan la decisión de incluir módulos narrativos en el robot desarrollado en esta tesis, entendiendo la narración como un mecanismo de enseñanza y aprendizaje afectivo mediado por tecnología.

Robots como mediadores en la comprensión de historias

El rol del robot en la educación infantil cambió: ya no se limita a narrar historias, sino que actúa también como mediador que ayuda a construir e interpretar su significado. Este segundo rol no solo se limita a narrar; también establece un diálogo con el niño que va más allá de los personajes, las acciones y las emociones que hay presentes en la historia. Dicha mediación hace

de la lectura una práctica de interacción colaborativa entre el niño y el robot en la que la comprensión se establece entre ambos.

Bizarro Torres et al. (2018) fueron de las primeras personas que exploraron esta idea, porque introdujeron el robot Roamer en las aulas infantiles para trabajar nociones espaciales en base a una narración guiada. Si bien el estudio no medía directamente el lenguaje, descubrió que el uso del robot en la activación de las narraciones mejoraba la atención de los niños y su comprensión secuencial, así como sus expresiones lingüísticas. Los escolares recordaban y relataban las acciones de la narración cuando el robot intervenía mediante preguntas y feedback con mayor coherencia desvelando una mediación robótica que provocaba una memoria narrativa, así como promover la estructuración lógica del relato, tal como también se presenta como una dimensión del lenguaje en desarrollo en la infancia (Bizarro Torres et al., 2018).

Continuando con la misma línea de investigación, Conti, Cirasa, Di Nuovo y Di Nuovo (2020) realizaron, de una forma parecida, el análisis sobre la posibilidad de que un robot social apoye la lectura compartida en el aula. En su análisis denominado, “Robot, ¡tell me a tale!", un robot humanoide leía los cuentos en compañía del docente, lanzaba preguntas abiertas y cambiaba de voz en función del tono que exigía cada parte de la historia. Los niños que participaron en el estudio tuvieron puntuaciones más altas en comprensión lectora y vocabulario que aquellos que recibieron únicamente la narración tradicional.

La teoría del andamiaje que nos ofrece Vygotsky tiene una interpretación de estos hallazgos: el robot funciona como un soporte que se interpone de forma temporal y permite el aprendizaje por medio de la participación aceptada. Al interrogar con preguntas abiertas y de forma recapituladora, favorece la consolidación de la comprensión del texto y la inferencia de algún tipo de significado.

Wang et al. (2025) continuaron en esta línea y exploraron las posibilidades de los robots narradores con inteligencia artificial generativa para el aprendizaje de la lengua en educación infantil. Su artículo "*Let Robots tell stories*" muestra cómo la integración de modelos de lenguaje avanzados permite a los robots ajustar la historia, elaborar preguntas a partir de las respuestas del niño e incluso modificar el discurso según esas respuestas. Dicha flexibilidad convierte el robot en un intermediario entre el contenido cognitivo y el emocional, de forma que puede reaccionar al estudiante, seguir su ritmo y proporcionar respuesta instantánea a sus pautas de trabajo. Tal como apuntan los resultados de Wang et al. (2025), hay significativos efectos sobre la comprensión inferencial de la historia, es decir, la habilidad de los niños para poder juzgar las intenciones, las causas y los efectos de la duración de la historia, esto es, su interpretación del texto. También aumentó la motivación y la atención mantenida, muy en especial cuando las expresiones del rostro y el adornos del robot eran congruentes con el contenido de la historia.

Estos estudios encuentran que la mediación robótica contribuye a la comprensión narrativa al mejorar la comprensión narrativa a través de la integración de elementos cognitivos, afectivos y sociales. En un sentido pedagógico, por ejemplo, el robot es una herramienta y facilitador en el aprendizaje dialógico, donde el niño juega para crear significado de la interacción. Si bien el maestro, de igual forma, sigue manteniendo un cierto liderazgo para la orientación global, la utilización del robot permite un espacio donde se logra dar lugar a la práctica individualizada y al descubrimiento individual mediante un lenguaje y preguntas adaptados al nivel de destreza del propio estudiante. Esta mediación se convierte, finalmente, en un ambiente de aprendizaje mucho más interesante y estimulante, y es necesaria en esos primeros años donde la atención y la lectura aún se encuentran en proceso de desarrollo.

Bizarro Torres et al. (2018), Conti et al. (2020) y Wang et al. (2025) aportan pruebas suficientes para concluir que los robots mediadores no sustituyen al docente, sino que amplían

las oportunidades de interacción para el alumnado. De esta manera, el diálogo, la retroalimentación y la ayuda emocional permiten un entendimiento en profundidad de las historias; refuerzan el pensamiento narrativo y fomentan la expresión verbal. Su incorporación en las aulas de preescolar de esta forma se plantea como un medio eficaz para la conjunción entre tecnología, desarrollo del lenguaje y la alfabetización inicial, objetivo primordial del presente proyecto.

Efectos de la emoción y la expresividad robótica en la motivación y aprendizaje

La implicación emocional en la interacción humano-robot se entiende como uno de los elementos que configuran las experiencias educativas mediadas por la tecnología en la medida en que las emociones afectan la capacidad para aprender, así como la atención, la memoria y la participación del niño. Esta es la razón por la cual la expressividad robótica, entendida como el grado en que un robot es capaz de expresar emociones a través de tono de voz, gestos o expresiones faciales, se ha convertido en una de las variables centrales a la hora de diseñar robots educativos. Más recientemente, se ha evidenciado que un comportamiento emocional coordinado por parte del robot resulta más atractivo para los niños y favorece tanto la cognición como el afecto en el proceso de aprendizaje (Appel et al., 2021).

El trabajo de Appel et al. (2021) proporciona un marco robusto para entender la forma en que la congruencia afectiva puede incidir en la interacción entre robots contadores de historias y niños. Según los autores, la concordancia entre el registro de los sentimientos del robot y el de la historia determina la inmersión narrativa y la imagen que se tiene del narrador: cuando existía esa conexión (una voz alegre si el pasaje era positivo, una voz melancólica en el caso de pasajes tristes), se mostraba empatía, y también se daba una mayor duración de la atención así como una mayor receptividad hacia la narración. Este resultado respalda la idea de que una emoción

coherente contribuye a la memoria y al conocimiento al activar, a la vez, una respuesta cognitiva y afectiva, y suma a la experiencia narrativa una dimensión humana y sensible.

De forma parecida, Ahmad et al. (2019) estudiaron cómo la retroalimentación emocional adaptativa incide en la motivación y en el aprendizaje de vocabulario. En su estudio longitudinal, el robot ajustaba sus respuestas emocionales al desempeño del niño durante las sesiones: se mostraba entusiasta cuando la respuesta era correcta y ofrecía gestos de aliento cuando no lo era. Esta adaptación mejoró la motivación intrínseca y la persistencia en la tarea, en particular entre los niños con mayor tendencia a distraerse. Según los autores, las funciones emocionales del robot no solo actúan como estímulo externo, sino como un mecanismo de regulación afectiva que ayuda al niño a sostener una actitud positiva hacia el aprendizaje, además de favorecer notablemente la retención de vocabulario gracias al vínculo emocional que se construye entre el niño y el robot.

Además, Blavette et al. (2025) demostraron la aceptabilidad y usabilidad de un robot social integrado con un modelo de lenguaje grande (LLM) para el cuidado geriátrico, pero su conclusión podría extrapolarse fácilmente al ámbito educativo dado los procesos de interacción socioemocional similares. Su estudio descubrió que los robots que son emocionalmente expresivos y coherentes parecen mostrar una emoción natural y coherente y tienen mayores niveles de empatía y confiabilidad, lo que genera una mayor disposición para la comunicación y colaboración por parte de los usuarios. En el contexto de la enseñanza infantil, se observa esto como una relación de confianza y conexión más cercana, condiciones necesarias para que los niños participen verdaderamente en las discusiones y cometan errores o expresen dudas en un proceso de apertura. La naturalidad en la comunicación emocional del robot es, de hecho, lo que dicta la aceptación y efectividad del robot como intervención pedagógica, no la calidad técnica, destacan los autores.

En definitiva, todos estos estudios corroboran de manera conjunta cómo la emoción o la expresividad robótica ejercen un efecto directo sobre la motivación, el compromiso y la efectividad del aprendizaje, ya que, desde el punto de vista pedagógico, el robot se transforma en un facilitador afectivo que en cierta manera se adapta a la experiencia que se desarrolla dependiendo del estado emocional del niño, al mismo tiempo que ayuda a mantener la atención y a reducir la ansiedad, especialmente en el caso de los más pequeños en edad preescolar, que es cuando la regulación emocional aún se está completando. Gestos, expresiones y tono de voz coherentes permiten, además, que los niños comprendan de manera natural las emociones del robot, y con eso profundizan tanto el lenguaje emocional como el verbal.

La expresividad emocional de los robots no es un adorno estético, sino un recurso pedagógico esencial que potencia la interacción significativa y la motivación por aprender. Las investigaciones de Appel et al. (2021), Ahmad et al. (2019) y Roštšintte et al. (2025) coinciden en que los robots más efectivos son aquellos capaces de responder afectivamente al contexto, generando experiencias educativas inmersivas y emocionalmente positivas. Estos principios se incorporan en el diseño del robot interactivo propuesto en la presente investigación, cuyo objetivo es integrar la emoción como un eje transversal en la enseñanza del lenguaje y la construcción de vínculos comunicativos auténticos con los niños.

Comparativa entre narración digital tradicional y narración con robots

Las estrategias narrativas en la educación infantil transitaron de los medios digitales unidireccionales hacia experiencias interactivas mediadas por robots sociales. La narración digital convencional (aplicaciones de narración, vídeos, plataformas multimedia) sigue siendo válida para la enseñanza de lenguas, pero con un rango limitado de interacción social y de adaptabilidad emocional. La narración asociada a robots incorpora en cambio una dimensión dialógica y emocional que transforma la narración en un proceso de ida y vuelta gracias a las

posibilidades del robot de responder, preguntar y manifestar en directo emociones (Essien & Parbanath, 2024).

Essien y Parbanath (2024) apuntan que el enfoque de la narración oral tradicional sería, en definitiva, exponer el contenido al niño a partir de la pantalla para que este lo aprenda por observación para, después, repetirlo; se trata de ambientes extremadamente estimulantes visual y auditivamente, pero que no otorgan la reciprocidad que es propia de la comunicación verbal entre los humanos. En su opinión, este formato favorece una comprensión "superficial" y un aprendizaje de vocabulario, pero no hace que se profundice el aprendizaje ni favorece la reflexión ni la producción oral del niño. Los robots narradores, en cambio, propician un aprendizaje conversacional en el que el niño participa, responde preguntas y construye un vínculo emocional con el narrador; esta personalización permite adaptar el contenido a las respuestas del usuario y a su nivel de comprensión, algo que los sistemas digitales tradicionales no logran ofrecer.

Papadopoulos et al. (2020) sostienen, en su revisión sistemática de robots sociales aplicados a la educación preuniversitaria, que la narración con robots aporta nuevos canales de comunicación, sino también una nueva forma de presencia pedagógica. En contraste con una pantalla, el robot ocupa un espacio físico dentro de la clase, lo cual produce una sensación de presencia compartida, de atención compartida que resulta propiciador del aprendizaje significativo. Este componente social le permite al niño conectar las narraciones con interacciones reales y auténticas -gestos, contacto ocular, turnos de la palabra-, relevantes en la construcción del lenguaje oral. También los autores subrayan que los robots pueden ajustar el timbre de voz y la tonalidad de sus gestos, de tal manera que puedan manifestar afecto y brindar así una interacción multisensorial, próxima a la comunicación humana, donde el aprendizaje ya no sólo se encuentra circunscrito a lo verbal, sino que también incluye lo no verbal y lo emocional.

Desde un punto de vista de la pedagogía, la diferencia básica entre la narración digital tradicional y la narración con robots reside en el hecho de que mientras la narración digital tradicional tiene un carácter predominantemente informativo, la narrativa digital con robots se basa en la interacción y en la construcción del conocimiento en forma colaborativa. En la narración digital tradicional el niño desempeña su papel, sobre todo, como receptor; en la narración digital con robots, en cambio, el niño actúa como un participante activo que, junto al robot, dará forma y sentido a la historia. Esta singularidad ubica la narración robótica en el marco de la construcción del conocimiento y de las perspectivas socioculturales, donde el aprendizaje se va produciendo a partir de la participación guiada y la interacción con agentes sociales: mantener un diálogo natural, tener una retroalimentación inmediata y la posibilidad de percibir reacciones emocionales del otro favorece tener un aprendizaje más motivado, una mejor retención del contenido y mayor capacidad de expresión oral.

Essien y Parbanath (2024) sostienen que personalizar los aspectos emocionales y lingüísticos del robot narrador hace que la narrativa robótica sea sustancialmente mejor que las plataformas digitales estáticas. Los actuales sistemas robóticos pueden si desean cambiar la dificultad del léxico, la entonación de la voz o la forma de realizar preguntas con arreglo al comportamiento o desempeño del estudiante y, así, ofrecer un entorno más completo y adaptable para el aprendizaje. La forma de narración digital tradicional, basada en un guion, no permite esa flexibilidad y no responde a las claves emocionales del niño; el robot, en cambio, interpreta su estado emocional, construyendo una relación educativa más personalizada y centrada en la persona.

Los costos, la complejidad técnica y la formación docente que exige incorporar la narración robótica son, no obstante, una advertencia que hacen Papadopoulos et al. (2020) pese a sus beneficios. Esto no quiere decir que se deban considerar los enfoques digitales y robóticos

como excluyentes, sino que, por el contrario, pueden ser complementarios dentro del mismo ecosistema educativo digital, este ecosistema puede combinar la sensorialidad propia de la narrativa digital con la interactividad y la empatía de la robótica social. Ambos formatos tienen el mismo objetivo de estimular el lenguaje y la imaginación de los niños, aunque con profundas diferencias entre ellos en cuanto a la propia experiencia comunicativa que proponen el primero centrado en el aprendizaje observacional y dirigido y el segundo centrado en la experiencia, el diálogo y la significatividad emocional.

La evidencia de Essien y Parbanath (2024) y Papadopoulos et al. (2020) apoya que la robótica narrativa es una evolución pedagógica natural de las tecnologías digitales hacia entornos más humanos, interactivos y adaptativos; estos son los principios que guían el presente proyecto.

Interacción humano-robot para el aprendizaje lingüístico

Esta sección se centra en la interacción humano-robot orientada al aprendizaje del lenguaje, y en cómo los avances de la robótica social han hecho posible construir modelos de comunicación más naturales, empáticos y adaptativos entre los niños y los robots educativos. Se revisan distintos enfoques de interacción que combinan cognición social, retroalimentación emocional y cooperación con el fin de reforzar la motivación, la atención y la adquisición del lenguaje en edades tempranas, junto con el papel que juegan la personalización y la adaptación emocional del robot para responder de forma sensible a las expresiones del niño y favorecer así un aprendizaje más significativo. Todo ello constituye la base conceptual sobre la que se apoya el diseño del sistema robótico propuesto en este trabajo.

Modelos de interacción niño-robot en contextos educativos

Hoy en día, la interacción social entre el niño y los robots sociales se ha extendido desde simples respuestas programadas hasta una serie de modelos de comunicación bidireccional que integran la cognición social, el lenguaje natural y la adaptabilidad emocional, que tienen como objetivo establecer dinámicas que sean características de la enseñanza humana, de la tutoría, de la cooperación, del aprendizaje entre pares, y que fomentan la motivación, la atención o el rendimiento de los más pequeños.

Bertacchini et al. (2023) presentan un modelo de interacción social constituido a partir de la interacción analizada a partir de la retroalimentación cognitiva y afectiva, apoyado en el modelo de lenguaje propuesto por el modelo ChatGPT encarnado en el robot. El robot logra, por eso, sostener una conversación activa con niños diagnosticados con Trastorno del Espectro Autista (TEA), al tiempo favorece el mantenimiento de atención conjunta, el vocabulario y la imitación verbal. El experimento demuestra la capacidad de adaptación conversacional: el robot adapta su habla y la entonación y el tono emocional de acuerdo con la respuesta que emite el niño, criterio clave para la perdurabilidad de la participación en el tiempo. Esta característica sobrepasa la ejecución estrictamente preprogramada y convierte al robot en un agente social empático, con la habilidad de adecuar su manera de comunicarse y reforzar los procesos cognitivos y del habla del niño.

Fang et al., 2025, también desarrollan paralelamente un modelo de colaboración humano-robot (HRC), pero orientado al desarrollo de la competencia socioemocional y el aprendizaje en equipos de preescolar. El robot, en este caso, no se sitúa en la figura del docente, sino en la figura de un compañero de equipo, que trabaja junto con los niños en las tareas cooperativas y la comunicación entre pares. Y los autores del trabajo identifican que esta forma de acercarse a la dinámica colaborativa mejoraba, medía y evaluaba la empatía, la regulación de las emociones

y el comportamiento prosocial además de la expresión oral y la comprensión de las instrucciones; así la propuesta muestra que los robots sociales pueden ayudar en el desarrollo del lenguaje no solamente mediante la enseñanza de lenguajes, sino e incluso en experiencias de trabajo de forma compartida que dejan espacio al lenguaje natural en las interacciones y en el proceso del aprendizaje social.

Roštšinskaja et al. (2025) exploraron el modelo de interacción escolar con el robot Pepper desde la usabilidad, la aceptación y la participación infantil. Basándose en observaciones realizadas en un conjunto de más de 100 niños y educadores, analizaron las respuestas en términos de atención y emociones ante la dinámica del robot pudiendo concluir que la aceptación del mismo depende en gran medida del antropomorfismo, de la sincronización multimodal de los gestos, la voz y la mirada y de la retroalimentación verbal y visual con buena percepción, concluyeron que el logro educativo del robot no sólo depende de la competencia técnica, sino también de la gestión de la confianza y del establecimiento de un vínculo social auténtico con el niño. Junto con los dos estudios anteriores, este último permite identificar tres tendencias complementarias en los modelos actuales de la interacción niño-robot:

1. La interacción adaptativa y emocional según la cual se verifica la personalización del discurso del robot y las emociones del mismo (Bertacchini et al., 2023).
2. La colaboración social y el aprendizaje colaborativo, en que el robot se comporta como amigo de juegos o de aprendizaje (Fang et al., 2025).
3. Interacción naturalista y multimodal, dirigida a optimizar la usabilidad y la inmersión social a través de gestos, voz y expresiones que sean congruentes unas con otras (Roštšinskaja, y otros, 2025).

Estos enfoques marcan un cambio de la instrucción unidireccional hacia una relación más recíproca, en la que el niño llega a percibir al robot como alguien con intenciones propias, capaz también de sentir. Este avance refleja lo que realmente representan los robots educativos: no se trata solo de transmitir información, sino de acompañar el desarrollo integral del lenguaje, la empatía y la interacción social. Los modelos de interacción descritos en esta sección constituyen, en ese sentido, la base conceptual del robot propuesto en esta tesis, que combinará narración, retroalimentación verbal y gestual, y detección de pronunciación para ofrecer una experiencia de aprendizaje más humana y significativa.

Personalización y adaptación emocional del robot

La capacidad que puede tener un robot social para modificar sus respuestas emocionales y su conducta, en función de las características que presenta cada niño, influye determinando, más que ninguna otra variable, el éxito de la interacción educativa. En el aula infantil, esa capacidad de modificar su conducta no solo incrementa la motivación y mantiene la atención por parte de los niños, sino que también contribuye a la construcción de un vínculo afectivo que posibilita la mejora en el aprendizaje del lenguaje y de la comunicación social. Según sugiere la literatura reciente, los robots que modulan su expresividad y su retroalimentación emocional logran mayor implicación cognitiva y afectiva por parte de los niños (Ahmad et al., 2019).

Un modelo de interacción a largo plazo, en el que un robot ofrecía retroalimentación emocional adaptativa durante sesiones de aprendizaje de vocabulario con niños en edad preescolar, fue el objeto de estudio de Ahmad et al. (2019). El sistema modulaba su respuesta emocional a partir de señales del niño —interés, entusiasmo, distracción o desgano— captadas mediante la expresión facial, el tono de voz y los gestos corporales. Los niños que interactuaron con el robot adaptativo sostuvieron la interacción social por más tiempo y aprendieron más palabras que quienes interactuaron con un robot de comportamiento fijo, lo que confirma el peso

de la sincronía emocional y la retroalimentación contingente para reforzar el vínculo entre afecto y aprendizaje, y para estimular una comunicación más espontánea.

Barry et al. (2025), a su vez, investigaron las diferencias individuales en la manera en que los niños interactúan con robots sociales según edad, temperamento y preferencia comunicativa. Hallaron que el estilo emocional del robot contrasta de una forma distinta en función del perfil del niño; algunos respondían mejor a un tono cálido y entusiasta, mientras que otros tenían más afinidad con un estilo más tranquilo y predecible. Los autores proponen que, a partir de ello, los robots educativos deberían orientar su respuesta a partir del monitoreo continuo del niño, variando el contenido emocional y la complejidad de sus intervenciones habladas; los sistemas más eficaces, apuntan, serían aquellos que sabrían aprender patrones de interacción que van surgiendo con cada niño y con ello argumentar una relación dinámica e orientada al usuario.

Ambos estudios coinciden en que la dimensión afectiva es indispensable para generar experiencias de aprendizaje significativas. Frente a los modelos tradicionales de enseñanza asistida por computadora, la robótica social añade una reciprocidad emocional que permite al niño no solo recibir información, sino también percibir una respuesta emocional coherente que ayuda a moldear su propio comportamiento comunicativo. La personalización emocional le permite al robot, en definitiva, ajustar su intervención según el ritmo y el estilo de aprendizaje de cada niño, su nivel de atención y su estado emocional, decidiendo incluso si conviene facilitar o hacer más exigente la tarea.

Con este tipo de estrategias, el robot no se limita a ejecutar una secuencia de lenguaje predefinida: también responde emocionalmente de forma coherente, ajustando voz y expresión facial a la situación, en línea con el enfoque de este proyecto. La idea es que el niño perciba al robot como un compañero de aprendizaje y no únicamente como un tutor funcional; una

herramienta empática que enseña a comunicarse tanto como a hablar, y que de ese modo despierta el interés del niño y lo mantiene comprometido con la actividad.

Tecnologías de reconocimiento y retroalimentación lingüística

En esta parte se examinan las innovaciones en tecnologías de reconocimiento, así como las de retroalimentación lingüística para que un robot educativo entienda, evalúe y genere un lenguaje que sea natural y adaptativo. Se repasan los avances en el reconocimiento automático del habla empleado por los niños (ASR) y las dificultades técnicas que conlleva, así como las estrategias de personalización que son necesarias para poder interpretar lo expresado por un niño, a la vez que se examina el efecto de la síntesis del habla (TTS) para establecer una comunicación que sea tanto expresiva como emocionalmente coherente, lo cual es muy importante en la retención de la atención del niño y para estrechar el vínculo emocional del niño durante las actividades. También se tratan en esta sección los mecanismos de retroalimentación fonética automatizada que permiten la corrección instantánea y empática de las fallas en la pronunciación y la integración multimodal de voz, imagen y gestos, permitiendo que la interacción se desarrolle en un contexto natural e inmersivo. Estos elementos constituyen, en conjunto, el marco operativo del sistema propuesto, orientado a favorecer la adquisición del lenguaje en un entorno cómodo, afectivo y pedagógicamente cercano.

Reconocimiento automático del habla infantil (ASR)

El reconocimiento automático del habla (ASR) es una de las tecnologías más relevantes para los robots educativos interactivos: permite interpretar y procesar de manera instantánea lo que dice el niño. El reto es mayor cuando se trata de voz infantil, que presenta variabilidad fonética, inestabilidad en la entonación, menor volumen y diferencias acústicas y anatómicas respecto a la voz adulta. Por eso buena parte de la investigación reciente se concentra en

construir modelos ASR robustos que puedan personalizarse para el habla infantil, un tema central en la robótica educativa y en el aprendizaje asistido por voz (Janssens et al., 2024).

Janssens et al. (2024) evaluaron la detección de voces infantiles en escenarios de interacción humano-robot utilizando modelos de vanguardia como Whisper y wav2vec 2.0. A pesar de las mejoras logradas en los últimos años, encontraron que en el aula real estos sistemas alcanzan una precisión cercana al 60%, debido tanto a la variabilidad acústica del habla infantil como a las interrupciones naturales que ocurren cuando los niños hablan. Los autores sugieren que avanzar en esta área requiere entrenar los modelos con corpus específicos de voz infantil y aplicar estrategias de ajuste dinámico que se adapten al hablante y al entorno; ese nivel de precisión técnica es, en última instancia, el que determina si el robot puede identificar correctamente las pronunciaciones y responder con las funciones educativas previstas.

Spaulding et al. (2018) propusieron, por su parte, un sistema robótico social para modelar la pronunciación infantil en tiempo real, apoyado en la API de SpeechAce para evaluar los fonemas emitidos por el niño. Esta situación permitía al robot ser capaz de repetir y comprobar la producción del habla, confrontándola con una referencia fonética, así como proporcionar también retroalimentación gestual y verbal. Por consiguiente, los niños se sintieron como si hubieran mejorado en autoeficacia comunicativa y en la autorregulación de su habla. Ellos mismos vieron de una manera positiva las correcciones que el robot ejerce sobre ellos, lo que les fue labrando poco a poco una distancia más producida en su producción articulatoria. Este trabajo revela cómo la interrelación entre la voz en reconocimiento y la retroalimentación social puede ser capaz de convertir la enseñanza de la pronunciación en un proceso interesante y gratificante.

Iio et al. (2024) realizaron una comparación entre los resultados proporcionados por tutores humanos y el sistema de aprendizaje que usa robots en el contexto de la enseñanza oral

del inglés como lengua extranjera. Su trabajo revela que, aunque los tutores humanos aún superan a los robots en las pruebas de habla, se dan resultados similares en otras variables como fluidez, repetición o pronunciación cuando el reconocimiento automático de voz (ASR) se utiliza en el contexto de los robots. Se han producido niveles de entusiasmo similares en los estudiantes y su ansiedad fue significativamente mayor a la de otros estudiantes que interactuaron usando los robots, lo que deja entrever que hay un efecto positivo en la motivación cuando se automatizan las interacciones. Estos datos sugieren que los modelos ASR pueden sustituir la instrucción personalizada en los casos donde los refuerzos continuos del habla no están disponibles por tiempo o por la ausencia de potencialidad de intervención humana.

En última instancia, los estudios recopilados evidencian el avance del ASR infantil hacia modelos más precisos, adaptativos e integrados socialmente con aplicaciones del mundo real en un entorno de aprendizaje. Aun así, todavía existen desafíos técnicos para detectar fonemas menos precisos, frases cortas y las expresiones no verbales que siguen al habla. Agregar un nuevo módulo ASR en el proyecto actual puede ayudar al robot a detectar respuestas verbales y errores de pronunciación cuando están contando historias durante la formación de oraciones. Este aspecto no solo mejorará la naturaleza interactiva del sistema, sino que también desarrollará una retroalimentación pedagógica inteligente individual para el nivel de desarrollo de cada niño.

Síntesis de voz en robots educativos

La síntesis de voz (TTS, Text-to-Speech) es uno de los pilares de la interacción humano-robot, sobre todo en entornos educativos donde la comunicación verbal es el medio principal de aprendizaje. En los niños en edad preescolar, la calidad, la naturalidad y la expresividad de la voz del robot influyen directamente en su comprensión, su atención y su vinculación emocional con el agente tecnológico. La investigación reciente se enfoca, por eso, en desarrollar sistemas

TTS capaces de generar voces más realistas, sincronizadas con la expresión facial y adecuadas a las emociones o intenciones comunicativas del contexto educativo (Esfandbod et al., 2023).

Esfandbod et al. (2023) llevaron a cabo el estudio de un robot emocional con dientes de precisión labio-timing en sesiones de terapia de habla para niños que presente trastornos del lenguaje. Dicho sistema combinaba un modelo de síntesis de voz expresiva que era capaz de modular la entonación y el ritmo de habla según el contenido emocional del mensaje con las acciones articulatorias labiales y faciales sincronizadas. Esta unión generó una comunicación más natural y con mayor credibilidad social y, por tanto, conducía a una mayor motivación de los niños y la obtención de una mejor imitación fonética de los niños durante la terapia. Sin embargo, los autores encontraron que la sincronía de tiempo del audio con el movimiento facial (en particular la apertura y el cierre de los labios) es fundamental para que el robot fuera percibido como un interlocutor humano.

La investigación también mostró que los métodos de síntesis de voz tradicionales en los sistemas TTS, que se basan en la concatenación de fonemas o bien en una modulación estándar, no aseguraban conseguir mantener a los niños interesados durante actividades de lenguaje que se extendían durante períodos de tiempo relativamente largos. En cambio, los modelos de TTS afectivos que modifican la prosodia, la velocidad o la intensidad de la voz de salida de acuerdo con la emoción y con la respuesta del usuario lograron una interacción más inmersiva y empática porque en este caso el robot expresa al mismo tiempo tanto el significado como también la emoción del mensaje, reforzando así tanto la comprensión semántica como el vínculo afectivo con el niño.

Esfandbod et al. subrayan, igualmente, la necesidad de una coherencia multimodal voz, gestos y mirada trabajando sinérgicamente, ya que los infantes tienden a establecer un vínculo entre la credibilidad del robot y la naturalidad en su comunicación no verbal en el momento

concreto. Este aspecto es especialmente importante en el caso de los robots educativos que cumplen una función como narradores o guías de tipo lingüístico, porque la variación en la entonación o la variación de las expresiones faciales puede favorecer la prosodia y la conciencia fonológica en los más pequeños.

La implementación de un sistema de síntesis de voz permitirá, en el presente proyecto, la expresividad emocional y comunicativa del robot, lo que facilita la narración de cuentos, la formulación de preguntas y la retroalimentación en actividades de pronunciación. Una conversión de texto a voz (TTS) que sea amigable, para hablantes de español, y se sincronice y combine con las animaciones de los movimientos faciales del robot permitirá que la interacción sea natural, comprensible y atrapante. La síntesis de voz no sólo es una herramienta técnica encargada de proveer el audio, sino que se convierte en un recurso de aprendizaje, que combina tecnologías con un aprendizaje significativo a través de lenguaje hablado.

Retroalimentación de pronunciación y corrección fonética automatizada

La retroalimentación automática que proporciona información sobre el uso correcto de la pronunciación constituye una de las características centrales de los sistemas de aprendizaje mediado por robots, sobre todo en cuanto a la enseñanza de las habilidades fonológicas y lingüísticas aplicadas a la educación de niños. Esta funcionalidad permite al robot identificar errores articulatorios, analizar la producción del habla y dar respuestas correctivas inmediatas, y genera así un ciclo de aprendizaje adaptativo en tiempo real. El valor pedagógico de este proceso está en que los niños reciben una corrección inmediata y personalizada sin experimentar la ansiedad o frustración que puede surgir ante la corrección humana directa (Spaulding et al., 2018).

Spaulding et al. (2018) implementó un sistema robótico orientado a la modelación y, al mismo tiempo, a la corrección de la pronunciación de los niños durante el proceso de aprendizaje

de vocabulario. Este robot se servía de un sistema de reconocimiento de la voz y cálculo fonético con los datos de una base de datos de pronunciaciones correctas. Por tanto, podía comparar los fonemas del niño con los de la base de datos y hacer una evaluación de la exactitud. Si había desviaciones, el robot podía ofrecer información al respecto a través de gestos, animaciones o expresiones verbales, teniendo en cuenta que era posible una promesa de corrección y motivación. Los niños respondieron con entusiasmo a estas sugerencias y repitieron las palabras con mayor precisión en sesiones posteriores, lo que muestra que la corrección fonética automatizada, presentada en un formato lúdico y social, aumenta tanto la atención como la motivación infantil, y a la vez refuerza la autoconciencia lingüística.

En una línea complementaria, Iio et al. (2024) compararon el desempeño de un sistema de aprendizaje de idiomas asistido por robot con el de tutores humanos, centrándose en las habilidades orales. Los robots que incorporan mecanismos de retroalimentación sonora lograron niveles de eficacia comparables a los de los maestros humanos en lo que respecta a la fluidez y la pronunciación. Esto se logró gracias a la evaluación del habla del usuario, realizada mediante algoritmos de reconocimiento de voz junto con un análisis espectrográfico, lo cual permite ofrecer comentarios personalizados según el tipo y la gravedad del error. Además, los autores notaron que los niños presentaban una actitud más relajada y colaborativa ante las correcciones del robot, lo que facilitaba una retroalimentación continua y no punitiva, un aspecto clave para mantener su confianza al comunicarse.

La efectividad de la corrección automática, según afirman ambos trabajos, no solo depende de la habilidad técnica del reconocimiento de voz: depende también de la forma social con la que se produzca dicha retroalimentación. Un robot que mezcla las respuestas acústicas con las no verbales —sonrisas, animaciones, refuerzos verbales positivos— proporciona una experiencia más parecida a la de las personas, en la cual el error pasa muy pronto a ser una

oportunidad de mejora inmediata y se favorece al mismo tiempo la consolidación fonológica y articulatoria del lenguaje infantil.

La incorporación de un módulo de retroalimentación fonética permitirá al robot valorar la pronunciación de las palabras o frases breves, detectar errores específicos de articulación y proporcionar correcciones adaptadas a cada niño y niña. El sistema no funcionará únicamente como narrador o interlocutor, sino que actuará también como asistente del lenguaje, es decir, participará en guiar la adquisición de la lengua por medio de la observación, la práctica repetitiva y la corrección empática, contribuyendo así al desarrollo de las competencias orales más precisas y a consolidar también el rol de robot como objeto educativo para fomentar el lenguaje en la primera infancia.

Integración multimodal: voz, visión y gestos para la interacción natural

La habilidad para combinar la articulación hablada mediante la utilización de la vista y de los gestos como canales de comunicación alternativos resulta esencial para que la interacción humano-robot pueda conceptualizarse como una interacción natural. En el contexto escolar infantil, esta habilidad permitirá al robot ofrecer una expresión oral de la información emitida, como también reconocer y expresar señales que propician la comunicación no verbal, de forma que su actuación pueda adecuarse a las dinámicas sociales de los humanos. Un diseño multimodal de este tipo eleva la intuitividad, la credibilidad y la empatía de la interacción, aspectos esenciales para sostener la atención y la participación de los niños en actividades de aprendizaje lingüístico (Sophokleous et al., 2021).

Sophokleous et al. (2021) proponen el uso de visión por computadora y análisis de gestos en plataformas de robótica educativa para construir un marco técnico que permita al robot percibir movimientos, expresiones faciales y posturas corporales de los estudiantes. Esta forma de procesamiento visual del sistema hace que este perciba señales de atención o emoción para

ajustar el contenido y el estilo de la enseñanza en tiempo real por ejemplo, según los autores Los sistemas multimodales mejoran la interacción educativa, ya que coordinan la voz, la expresión facial y los gesto, refuerzan el mensaje, mostrándose que un robot que es capaz de combinar voz y gesto es capaz de tener un mayor aprendizaje de vocabulario en los niños y mayor participación que el del robot que sólo actúa verbalmente.

Fink (2012) argumenta, en cambio, que el antropomorfismo y la semejanza humana propios de los robots sociales no deben entenderse solo como resultado de su diseño mecánico, sino también de la coherencia comunicativa entre sus distintos canales. Su planteamiento sostiene que la interacción natural genuina surge cuando el robot logra sincronizar las modalidades vocal, visual y motora, y genera así intenciones y presencia social. Esa congruencia entre canales lleva a los usuarios, en especial a los niños, a atribuirle al robot rasgos humanos como sentimientos o comprensión, y con eso fortalece la confianza y la participación en el aprendizaje. No obstante, Fink advierte que las contradicciones entre voz y movimiento —"habla feliz vs. movimientos rígidos", por ejemplo— destruyen la coherencia en las emociones y destruyen la credibilidad del robot como recurso educacional.

Y la evidencia que surge de Sophokleous et al. (2021), junto con la de Fink (2012), es que la sincronía comunicativa es esencial para conseguir una experiencia educacional rica y emocionalmente valiosa. Un robot que habla, escucha y gesticula coordinadamente establece una relación más verosímil con el niño, lo que permite que se produzca tanto el aprendizaje lingüístico como las habilidades sociales y emocionales; es más, es capaz de detectar señales de implicación o de distracción y de modificar en consecuencia su conducta pedagógica, generando así una capa adicional en la sensibilidad al contexto.

Por tanto, en el presente trabajo confluirán de forma conjunta la voz, la visión y los gestos, con lo que el robot puede llegar a actuar con todas las de la ley como interlocutor: narrar historias,

hacer movimientos expresivos o dar respuestas visuales a la conducta del niño/a. Esta confluencia refuerza la autenticidad comunicativa y propicia el aprendizaje de la lengua al activar estímulos de auditivos, visuales y cinestésicos, lo que la sostiene y favorece la comprensión semántica. La multimodalidad se convierte así en un eje del diseño, orientado a construir un aprendizaje fluido, emocional y pedagógicamente significativo.

Revisión de estudios y análisis de tendencias

Esta sección revisa los estudios y las tendencias en robótica social aplicada a la educación infantil, con el fin de identificar los principales avances, enfoques metodológicos y vacíos presentes en la literatura. Presenta revisiones sistemáticas sobre el impacto pedagógico y emocional de los robots sociales en el aprendizaje temprano, junto con análisis bibliométricos que muestran la evolución, la distribución y los temas emergentes de esta investigación a nivel mundial. Discute, también, hacia dónde apunta la innovación a partir de las limitaciones detectadas, en particular en lo referido a robots lingüísticos adaptativos que integren narración, reconocimiento de voz y retroalimentación fonética para el aprendizaje del lenguaje en la primera infancia.

Revisión sistemática de la robótica social en educación infantil

La robótica social en la educación infantil se establece como un área interdisciplinaria que integra ingeniería, psicología del desarrollo y pedagogía para mejorar el aprendizaje mediante la interacción con robots. En los últimos veinte años, distintos estudios han mostrado que los robots sociales cumplen diversos roles —tutores, compañeros de juego, facilitadores emocionales— orientados a fortalecer habilidades cognitivas, sociales y lingüísticas. Para dimensionar el alcance de esta tecnología conviene revisar críticamente el estado actual del campo a partir de fuentes analizadas sistemáticamente (Papadopoulos et al., 2020).

Papadopoulos et al. (2020) realizaron una exhaustiva revisión de la literatura sobre robots sociales en educación, desde 2005 hasta 2020. El trabajo que presentan muestra que la robótica social se usa notablemente en educación infantil y primaria para fomentar las habilidades socioemocionales, comunicativas y de alfabetización inicial. Los robots que se emplean más a menudo son NAO, Pepper y KASPAR porque son capaces de reproducir emociones y mantener conversaciones elementales, lo que facilita la participación de los niños en actividades de tipo cooperativo o de narración en las que resulta relevante el que el robot interaccione de manera apropiada. La revisión también considera tres dominios que servirán de base para justificar la efectividad pedagógica de estos robots: el diseño del comportamiento social, la calidad de la interacción multimodal y la integración coherente de estos robots en estrategias didácticas estructuradas. Los autores llegaron a la conclusión de que la robótica social no reemplaza al docente, sino que representa una fuerza neutral que puede favorecer el aprendizaje o bien puede enriquecer el clima emocional del aula.

En la misma línea, Zorrilla-Puerto et al. (2023) llevaron a cabo una revisión sistemática donde el foco principal del estudio se centró en la robótica educativa durante la etapa infantil, buscando centrar su impacto del desarrollo de competencias del tipo cognitivo y socioemocional. Estudiaron publicaciones de bases de datos internacionales entre los años 2010 a 2022, seguidamente clasificaron los artículos y publicaciones dependiendo del tipo de robot, de la metodología empleada y de los resultados obtenidos y concluyeron que la robótica educativa mejora el desarrollo del pensamiento lógico, la ejecución de problemas y la comunicación oral hecho que se produce, eso sí, siempre que haya mediación de la figura adulta. Indican, asimismo, que los robots sociales colaboran en potenciar el trabajo en grupo de los niños, así como la empatía hacia los otros; aunque también reconocen que la mayor parte de los proyectos revisados carecen de evaluaciones a largo plazo y de métricas estandarizadas que permitan asegurar unos resultados sobre la efectividad educativa razonables.

Las dos revisiones coinciden en que la robótica social para la educación preescolar ha pasado de ser un simple tipo de aplicación experimental a ser una herramienta de aprendizaje emergente con un potencial innegable. La evidencia disponible sugiere que los robots pueden elevar la motivación, la comprensión verbal y la alfabetización temprana, sobre todo cuando se integran en contextos narrativos o lúdicos, aunque también se subraya la necesidad de avanzar hacia modelos más personalizados, receptivos y emocionalmente congruentes, capaces de ajustar su comportamiento al progreso y al estado emocional de cada niño.

Los resultados de estas dos revisiones respaldan, dentro del alcance de este proyecto, la necesidad de un robot lingüístico interactivo que integre narración de cuentos, reconocimiento de voz y retroalimentación fonética, y que combine una interacción social significativa con principios de aprendizaje multimodal. La propuesta desarrollada en esta tesis se alinea, en ese sentido, con los desarrollos más recientes de la literatura científica y constituye un aporte práctico a la robótica educativa aplicada a la enseñanza de idiomas en el aula infantil.

Análisis bibliométrico de la investigación sobre robots sociales

El análisis bibliométrico es una herramienta clave para entender la evolución, las tendencias y el alcance global de la investigación en robótica social. A partir de indicadores como el número de publicaciones, las redes de coautoría, las citas y las instituciones más influyentes, permite identificar temas emergentes, áreas de mayor impacto y vacíos en la literatura. En el campo educativo resulta especialmente útil, pues ofrece una mirada estructurada sobre cómo ha avanzado, desde lo tecnológico y lo pedagógico, la aplicación de robots como apoyo a la educación infantil (Zhang et al., 2024).

Zhang et al. (2024) llevaron a cabo un análisis que consiste en bibliometría y la visualización de datos de un periodo de más de 20 años de publicaciones dedicadas al uso de robots sociales para el aprendizaje y el desarrollo infantil, indexados en la Web of Science y

Scopus, observando que era un área de investigación en crecimiento, a partir de 2010 pero que da un salto evidente a partir de 2018 que tiene correlato con el desarrollo de la inteligencia artificial, el procesamiento de lenguaje natural y el reconocimiento de voz. Esta investigación clasifica los ejes de investigación en tres líneas de investigación principales:

1. Interacción Interacción social y emocional con robots por la educación
2. Estudio de evaluación del impacto cognitivo y del impacto lingüístico
3. Diseño de plataformas robóticas adaptativas para entornos educativos.

Los mismos autores comentan que el enfoque investigativo se centra, sobre todo, en Estados Unidos, Japón, Corea del Sur y algunos de los países miembros de la UE, lo cual evidencia un sesgo geográfico en la producción científica: escasa presencia de estudios producidos a partir de América Latina y de África.

Wang et al. (2025) realizaron, en cambio, un análisis de redes de colaboración temática e internacional en la investigación de robótica social para el desarrollo infantil, apoyado en modelado de temas y análisis de co-ocurrencia de palabras clave. Su estudio identifica temas emergentes como la inteligencia artificial, la narración digital asistida por robots y la integración de modelos de lenguaje grande en la educación, y describe una tendencia hacia la hibridación entre robótica y aprendizaje automático, en la que el robot deja de ser una simple entidad física para funcionar como una unidad cognitiva capaz de aprender del entorno y adaptar su comportamiento pedagógico. El análisis confirma, además, que la cooperación interdisciplinaria entre ingeniería, neurociencia y educación elevó la calidad y el alcance de las publicaciones en este campo.

Al comparar ambos análisis se aprecia que la investigación en robótica social gana madurez y especialización con el tiempo. Zhang et al. (2024) documentan un crecimiento cuantitativo de la producción científica y la consolidación de la comunidad académica, mientras

que Wang et al. (2025) ofrecen una lectura más cualitativa de los temas emergentes y las colaboraciones interdisciplinarias. Ambos estudios muestran, en conjunto, que el interés académico superó la mera validación técnica de prototipos para incorporar evaluaciones de impacto educativo, social y afectivo, con un énfasis creciente en procesos centrados en el niño.

Para los fines de esta tesis, un robot social educativo centrado en habilidades lingüísticas resulta pertinente, dado que el lenguaje y la comunicación son campos particularmente activos y productivos dentro de esta línea de investigación. Por otro lado, evidencian la necesidad de sólidos estudios en contextos hispanohablantes, contribuyendo, así, a diversificar una producción científica que continúa siendo producto de pocas regiones, a través de un trabajo adaptado a las particularidades lingüísticas, culturales y pedagógicas de América Latina.

Brechas y oportunidades para el diseño de robots lingüísticos

A pesar de todas las mejoras que se han dejado ver, la revisión de la literatura da cuenta de la existencia de brechas conceptuales, técnicas y metodológicas que constituyen un lastre para consolidar sistemas realmente integrales para el aprendizaje temprano de la lengua. A pesar de que muchos documentos científicos informan que los robots sociales mejoran la motivación, la interacción emocional y la adquisición de vocabulario, existe una disyunción técnica importante entre las partes que sustentan la comunicación natural, como el reconocimiento de habla natural humana, el análisis fonético y los motores de narración adaptativos que continúan desarrollándose, en buena medida, de forma separada.

Por una parte, la mayoría de los prototipos educativos recolocados en el estado de la cuestión se centran el manejo de tareas concretas de narración de cuentos o la práctica del vocabulario sin tener en cuenta mecanismos de ASR (reconocimiento automático del habla) capaces de realizar la precisión de la pronunciación infantil. Como señalan Janssens et al. (2024), los sistemas actuales presentan un margen de error considerable frente a la variabilidad

acústica del habla de los niños, lo que limita su capacidad de ofrecer retroalimentación inmediata y personalizada; esto ha llevado a que buena parte de las interacciones se apoyen en comandos predefinidos o respuestas de opción múltiple, en lugar de conversaciones libres.

El análisis fonético automatizado, en segundo lugar, no ha recibido mucha atención en la educación preescolar. Estudios como los de Iio et al. (2024) y Spaulding et al. (2018) demostraron la viabilidad de sistemas de corrección fonética mediante algoritmos de comparación de pronunciación, pero estos modelos todavía no se integran con facilidad a módulos emocionales y narrativos. La retroalimentación se reduce, en su mayor parte, a un juicio binario ("correcto/incorrecto"), que no toma en cuenta las dimensiones pedagógicas de dar refuerzo positivo o empatía, importantes para que un niño se mantenga motivado y competente en la práctica del habla.

La narrativa adaptativa —conceptualizada como 'cambiar las pautas de narrar de acuerdo a la respuesta del usuario'— es, en relación con la anterior, una funcionalidad que sigue siendo poco común en la literatura. La propuesta de Ahmad et al. (2019) o las de Bertacchini et al. (2023) pueden incorporar cierto grado de respuestas emocionales y adaptación al lenguaje, pero carecemos aún de una propuesta robusta que permita integrar el reconocimiento de voz infantil con la narración de historias a partir de narraciones o respuestas completas. Esta falta de integración impide que el robot funcione como un mediador lingüístico inteligente que ajuste sus decisiones narrativas y fonéticas a las necesidades de cada niño.

La suma de estas limitaciones pone de manifiesto que se encuentra un desarrollo todavía incipiente de sistemas robóticos multimodales y adaptativos con fines de aprendizaje lingüístico integral. La investigación llevada a cabo se ha centrado, fundamentalmente, en la evaluación de la tecnología o de la interacción social por separado, pero no tanto en la vinculación de la percepción —voz, visión— con las estrategias pedagógicas que fortalezcan el lenguaje. A ello

se une la inexistencia de estudios longitudinales que midan el rendimiento del uso de robots en el aprendizaje del lenguaje oral que se alargue en el tiempo, sobre todo en el caso de una población hispanohablante y de primera infancia.

Este estado de cosas nos abre una puerta evidente a la innovación: un robot que combine la narración interactiva, el análisis fonético automatizado y el reconocimiento de habla infantil podría dar lugar a una forma de aprendizaje más orgánica y empática. Dicha integración permitiría a este robot no sólo contar historias y formular preguntas, sino también evaluar pronunciaciones, determinar errores de articulación y empatizar con retroalimentación emocionalmente adaptada. Sumar tecnología de seguimiento facial y gestual mediante visión artificial podría, además, mejorar la sincronización entre la atención visual y el habla, y elevar así la calidad general de la interacción.

Los robots lingüísticos adaptativos representan, en definitiva, un campo de investigación todavía emergente, que surge de la confluencia entre inteligencia artificial, robótica social y enseñanza de idiomas. Esta tesis busca aportar a llenar ese vacío con el desarrollo de un robot interactivo que integre reconocimiento de voz, análisis fonético y narración adaptativa, orientado a apoyar el aprendizaje del lenguaje oral en niños de 4 a 5 años, y contribuir así tanto al avance tecnológico como al fortalecimiento de modelos pedagógicos basados en la comunicación social y el aprendizaje significativo asistido por robots.

A continuación, se presenta la Tabla 1, en la cual se sintetizan los principales hallazgos obtenidos durante la revisión bibliográfica realizada para el estado del arte.

Tabla 1
Resumen de Trabajos consultados

Autores	Año	Título del trabajo	Principales Hallazgos

Ahmad, M. I., Mubin, O., Shahid, S., & Orlando, J.	2019	Robot's adaptive emotional feedback sustains children's social engagement and promotes their vocabulary learning	Retroalimentación emocional adaptativa mejora la motivación, atención y aprendizaje de vocabulario en preescolares.
Appel, M., Lugrin, B., Kühle, M., & Heindl, C.	2021	*The emotional robotic storyteller: On the influence of affect congruency on narrative transportation, robot perception, and persuasion.*	El tono y emoción del robot aumenta la inmersión narrativa, la percepción positiva del robot y la persuasión.
Barry, R.-J., Neumann, M., & Neumann, D.	2025	Individual differences and young children's engagement with a social robot	A cada niño se le observan preferencias individuales en estilos emocionales del robot; se cita que la personalización del tono y las respuestas son necesarias.
Bertacchini, F., Demarco, F., Scuro, C., Pantano, P., & Bilotta, E.	2023	A Social Robot Connected with ChatGPT to Improve Cognitive Functioning in ASD Subjects	Conexión de robot con ChatGPT mejora la atención y vocabulario en niños con TEA; demuestra el valor de la interacción adaptativa y emocional.
Bizarro Torres, N., Luengo González, R., & Carvalho, J. L.	2018	Roamer, un robot en el aula de Educación Infantil para el desarrollo de nociones espaciales básicas	Roamer facilita el aprendizaje de nociones espaciales en preescolar mediante interacción lúdica; mejora atención y expresión verbal en tareas secuenciales.
Blavette, L., Dacunha, S., Alameda-Pineda, X., Hernández García, D., et al.	2025	*Acceptability and Usability of a Socially Assistive Robot Integrated With a Large Language Model*	La integración de modelos de lenguaje mejora la naturalidad y empatía percibida del robot; la expresividad emocional aumenta confianza y comunicación.
Breazeal, C., Dautenhahn, K., & Kanda, T.	2016	*Social Robotics* (Capítulo en Springer Handbook of Robotics)	Robots sociales como agentes que facilitan interacción, motivación y aprendizaje, importancia de diseño social en contextos educativos.
Chatzichristofis, S. A.	2023	Recent Advances in Educational Robotics	Tendencias a integración de Inteligencia artificial, visión y multimodalidad; destaca impacto en aprendizaje activo y personalización.

Conti, D., Cirasa, C., Di Nuovo, S., & Di Nuovo, A.	2020	Robot, tell me a tale!: A social robot as tool for teachers in kindergarten	Utilizar a un robot como instrumento docente mejora la habilidad de comprensión lectora y el vocabulario que muestra la mediación docente (andamiaje y preguntas abiertas).
Esfandbod, A., Rokhi, Z., Meghdari, A. F., Taheri, A., Alemi, M., & Karimi, M.	2023	Utilización de un robot emocional capaz de sincronización labial en sesiones de terapia del habla	Una voz sintética que se ejecute con el propósito de ser expresiva y sincronizada mejora la motivación para aprender e imita mejor la fonética; una coherencia audiovisual imprime mayor naturalidad en la comunicación.
Essien, G., & Parbanath, S.	2024	Exploring the World of Robot-Assisted Digital Storytelling: Trends, Models, and Educational Implications	Clasifica modelos (robots autónomos, modelos colaborativos y modelos híbridos); la conclusión que extraen, es la de una mejoría de los enfoques colaborativos y multimodales para una mayor participación y para la alfabetización.
Fang, J.-W., Ji, T.-T., Tu, Y.-F., Hwang, G.-J., Zou, D., & Chen, J.	2025	Effects of a human-robot collaborative teaching approach on preschoolers' social-emotional competence through collaboration between the child and robot	Modelo de trabajo conjunto humano-robot promueve la empatía, la regulación emocional y las habilidades lingüísticas mediante la cooperación social.
Feldman, H. M.	2019	The Importance of Language-Learning Environments to Child Language Outcomes	Resalta que entorno lingüístico y retroalimentación son críticos para resultados de lenguaje; contexto y calidad de interacción influyen más que solo cantidad de exposición
Fink, J.	2012	Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction	La combinación del realismo humano junto con la coherencia multimodal mejora la confianza, la empatía y la credibilidad educativa del robot.
Chen, Park, & Breazeal	2020	Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement	El aprendizaje recíproco con un robot aumenta el compromiso emotivo y mejora el aprendizaje.

Iio, T., Yoshikawa, Y., Ogawa, K., & Ishiguro, H.	2024	Comparison of Outcomes Between Robot-Assisted Language Learning System and Human Tutors	Los robots logran una eficacia comparable a tutores humanos en habla, favorecen la práctica fluida y reducen ansiedad comunicativa.
Janssens, R., Verhelst, E., Abbo, G. A., et al	2024	Child Speech Recognition in Human-Robot Interaction: Problem Solved?	Los modelos con reconocimiento automático de voz logran un 60% de precisión en voz infantil.
Kory, J. M., & Breazeal, C.	2014	*Storytelling with Robots: Learning Companions for Preschool Children's Language Development* (RO-MAN)	Los robots narradores mejoran la adquisición de vocabulario y estructuras narrativas fomentando la emulación lingüística.
Kory-Westlund, J. M., & Breazeal, C.	2019	A Long-Term Study of Young Children's Rapport, Social Emulation, and Language Learning With a Peer-Like Robot Playmate in Preschool	Una relación sostenida con un robot similar a un compañero mejora la confianza, la imitación social y el aprendizaje del lenguaje a lo largo del tiempo.
Wang, X., Li, L., Wang, Q., Zhong, B., & Xu, Y.	2023	*Meta–analyzing the Impacts of Social Robots for Children's Language Development*	Se documentan efectos generales positivos en lenguaje, variar dependiendo de la metodología utilizada y de la edad; ser necesarios estudios más estandarizados.
Mazı, A., & Yıldırım, İ. O.	2025	Primary school teachers' opinions on the use of artificial intelligence in educational practices	Los docentes muestran interés, pero también preocupaciones, la aceptación condiciona la facilidad de uso y soporte pedagógico.
Neumann, M. M.	2020	Social Robots and Young Children's Early Language and Literacy Learning	Los robots contribuyen al desarrollo del lenguaje y la alfabetización emergente
Papadopoulos, I., Lazzarino, R., Miah, S., Weaver, T., et al.	2020	A systematic review of the literature regarding socially assistive robots in pre-tertiary education	Los robots sociales favorecen el aprendizaje interactivo y emocional, pero hace falta más evidencia en el ámbito de la educación inicial.
Pasalidou, C., Lytridis, C., Tsinakos, A., & Fachantidis, N.	2025	Augmented Reality and Robotics in Education: A Systematic Review	La realidad aumentada y la robótica mejoran la atención y la comprensión al combinar actividades físicas y digitales. Se recomienda integrarlas en la

			enseñanza y capacitar a los docentes.
Roštšinskaja, A., Saard, M., Korts, L., Kööp, C., et al.	2025	Unlocking the Potential of Social Robot Pepper: A Comprehensive Evaluation of Child-Robot Interaction	Aceptación infantil depende del antropomorfismo, gestos coherentes y sincronía multimodal; resalta la importancia del vínculo social en la eficacia educativa.
Sariyati & Wardhani, J. D.	2023	*Improving Language Skills Through Storytelling Method with Audio Visual Media* (ICOLAE)	Mediante el storytelling con medios audiovisuales se produce una mejora de las habilidades orales en la educación infantil y sugiere interaccionar con los mismos para conseguir mejores resultados.
Smakman, M., Vogt, P., & Konijn, E. A.	2021	Moral considerations on social robots in education: A multi-stakeholder perspective	A continuación discute las implicaciones éticas (privacidad, dependencia, responsabilidad) y la necesidad de políticas/guías para implementar robots, a la hora de utilizarlos en las aulas.
Sophokleous, A., Christodoulou, P., Doitsidis, L., & Chatzichristofis, S. A.	2021	Computer Vision Meets Educational Robotics	Integración de visión y gestos permite detectar atención y emociones y que la enseñanza sea en vivo.
Spaulding, S., Chen, H., Ali, S., Kulinski, M., & Breazeal, C.	2018	A Social Robot System for Modeling Children's Word Pronunciation	Robot con reconocimiento de voz y corrección fonética mejora pronunciación infantil mediante retroalimentación inmediata y lúdica.
Wang, X., Wang, Y., Han, Z., Wang, Y., Duan, Z., Zhang, Z., & Alsudais, T. A.	2025	Social Robots for Child Development: Research Hotspots, Topic Modeling, and Collaborations	Destaca la interdisciplinariedad e innovación en educación infantil.
Wang, Z., Law, T. S.-T., & Yeung, S. S.	2025	Let robots tell stories: Using social robots as storytellers to promote language learning among young children	Robots narradores adaptativos que hacen uso de IA narrativa mejoran comprensión inferencial y motivación.

Zhang, N., Xu, J., Zhang, X., Wang, Y., et al.	2024	Social Robots Supporting Children's Learning and Development: Bibliometric and Visual Analysis	Emoción, lenguaje y aprendizaje adaptativa. Resaltan la división entre los investigadores en esta area
Zorrilla-Puerto, J., Lores-Gómez, B., Martínez-Requejo, S., & Ruiz-Lázaro, J.	2023	El papel de la robótica en Educación Infantil: revisión sistemática para el desarrollo de habilidades	La robótica mejora lenguaje y habilidades cognitivas, pero escasean estudios sobre pronunciación y fonética.

Capítulo III

En este capítulo se plantea el diseño conceptual del robot pedagógico propuesto y se narra la manera en la que se traducen las necesidades del usuario en parámetros técnicos y decisiones de ingeniería. Se presentan las etapas de identificación de requerimientos, análisis a partir de la matriz QFD, evaluación de alternativas y elección del concepto final que orientará el prototipo, se describen los módulos principales que conforman el sistema que son: materiales, estructura mecánica, actuadores, visualización, control y alimentación y se argumenta la elección técnica de cada uno de ellos. Con esto, el capítulo sienta la base metodológica y argumentativa necesaria para comprender la configuración final del robot y su correspondencia con los objetivos pedagógicos planteados.

Diseño Conceptual

A continuación se expone el proceso completo de selección de diseños, que comprende el planteamiento de las necesidades, las propuestas de diseño y la selección final. En la Tabla 2 se presentan las necesidades que el robot debe satisfacer.

Tabla 2
Necesidades del robot

Necesidades	Interpretación
Aspecto amigable.	Para fomentar la interacción y la conexión emocional con los niños, el diseño del robot debe ser atractivo y familiar, utilizando formas y colores que generen confianza y simpatía.
Brazos y cabeza que puedan simular gestos.	Relacionado con el aspecto amigable, los brazos y la cabeza deben tener movilidad para mantener la atención del infante
Pantalla para visualización gráfica de tamaño adecuado	Debe contar con una pantalla de 10” a color

Permitir la interacción táctil con las aplicaciones en pantalla	La pantalla debe tener un sistema táctil
HMI interactivo y sencillo	Un sistema de interfaz HMI intuitivo y claro permite que los niños naveguen con facilidad, se debe evitar la saturación de información.
Emitir sonidos y dar indicaciones	Un parlante que permita emitir sonidos básicos.
Escuchar las respuestas del usuario	Un micrófono que permita que el robot reconozca las respuestas orales de los niños
Fácil de transportar	Se debe evitar la presencia de piezas sueltas o flojas en el robot para facilitar su transporte.
De fácil montaje	No debe tener un exceso de piezas que dificulten su montaje.
Juegos de corrección de pronunciación y cuentacuentos	Solicitud especifica de las aplicaciones que deben ser programadas.

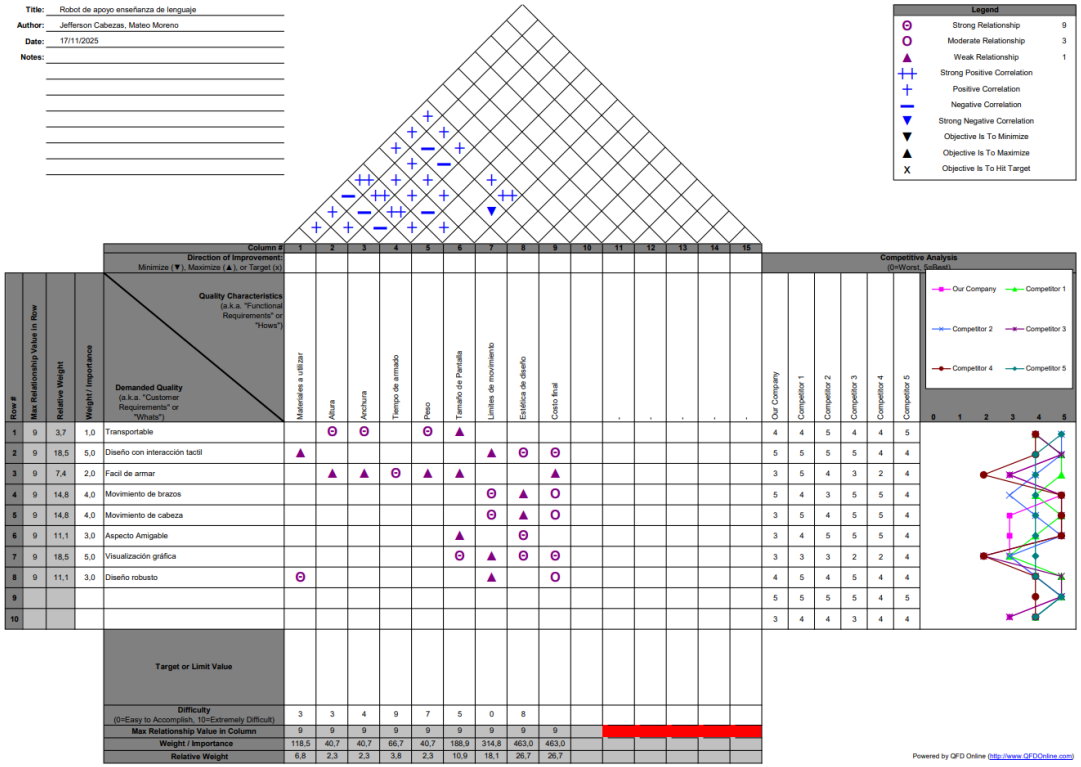
Nota. Elaboración propia.

El robot solicitado de manera general es un asistente pedagógico infantil con diseño amigable y expresivo, que incluye cabeza y brazos móviles, pantalla táctil de 10” con interfaz HMI sencilla, micrófono y parlante para interacción por voz, y aplicaciones educativas como cuentacuentos y juegos de pronunciación. Debe ser fácil de montar y transportar.

Matriz QFD

Definidos los parámetros de los requerimientos se procede a realizar la matriz QFD (Quality Function Deployment) que se muestra en la Figura 5, con el fin de convertir las características solicitadas por el cliente en parámetros técnicos de diseño.

Figura 5
Matriz QFD



Title: Robot de apoyo enseñanza de lenguaje
Author: Jefferson Cabezas, Mateo Moreno
Date: 17/11/2025
Notes:

Legend
Strong Relationship 9
Moderate Relationship 3
Weak Relationship 1
Strong Positive Correlation
Positive Correlation
Negative Correlation
Strong Negative Correlation
Objective Is To Minimize
Objective Is To Maximize
Objective Is To Hit Target

Row #	Max Relationship Value in Row	Relative Weight	Weight / Importance	Demanded Quality (a.k.a. "Customer Requirements" or "Whats")	Materiales a utilizar	Altura	Anchura	Tiempo de armado	Peso	Tamaño de Pantalla	Límites de movimiento	Estética de diseño	Costo final	Our Company	Competitor 1	Competitor 2	Competitor 3	Competitor 4	Competitor 5
Column #					1	2	3	4	5	6	7	8	9						
Direction of Improvement:				Minimize (▼), Maximize (▲), or Target (x)															
1	9	3,7	1,0	Transportable		⊖	⊖		⊖	▲				4	4	5	4	4	5
2	9	18,5	5,0	Diseño con interacción tactil	▲						▲	⊖	⊖	5	5	5	5	4	4
3	9	7,4	2,0	Facil de armar		▲	▲	⊖	▲	▲			▲	3	5	4	3	2	4
4	9	14,8	4,0	Movimiento de brazos							⊖	▲	O	5	4	3	5	5	4
5	9	14,8	4,0	Movimiento de cabeza							⊖	▲	O	3	5	4	5	5	4
6	9	11,1	3,0	Aspecto Amigable						▲		⊖		3	4	5	5	5	4
7	9	18,5	5,0	Visualización gráfica						⊖	▲	⊖	⊖	3	3	3	2	2	4
8	9	11,1	3,0	Diseño robusto	⊖						▲		O	4	5	4	5	4	4
9														5	5	5	5	4	5
10														3	4	4	3	4	4
Target or Limit Value																			
Difficulty (0=Easy to Accomplish, 10=Extremely Difficult)					3	3	4	9	7	5	0	8							
Max Relationship Value in Column					9	9	9	9	9	9	9	9	9						
Weight / Importance					118,5	40,7	40,7	66,7	40,7	188,9	314,8	463,0	463,0						
Relative Weight					6,8	2,3	2,3	3,8	2,3	10,9	18,1	26,7	26,7						

Competitive Analysis (0=Worst, 5=Best)

Our Company, Competitor 1, Competitor 2, Competitor 3, Competitor 4, Competitor 5

0 1 2 3 4 5

Powered by QFD Online (http://www.QFDOnline.com)

Nota. Elaboración propia.

Con los resultados obtenidos en la matriz QFD se muestran los resultados según su porcentaje de aprobación de la Tabla 3.

Tabla 3
Necesidades del robot

Característica	Porcentaje de Aprobación
Visualización gráfica	18,5
Interacción Táctil	18,5
Movimiento de brazos	14,8
Movimiento de Cabeza	14,8
Aspecto Amigable	11,1
Diseño robusto	11,1
Fácil de Armar	7,4
Transportable	3,7

Nota. Elaboración propia.

Se observa que las características primordiales son la visualización gráfica a través de una pantalla con interacción táctil, seguido del movimiento de los componentes del robot, y el aspecto del mismo, la facilidad de armar y la transportabilidad quedan como últimas necesidades en la lista.

Con estos resultados se definen los siguientes módulos para su análisis individual: material general en la Tabla 4, material para impresión 3D en la Tabla 5, tarjeta de control en la Tabla 6, tarjeta de control secundaria en la Tabla 7, visualización en la Tabla 8, modelo de pantalla en la Tabla 9, tecnología para el movimiento de brazo en la

Tabla 10, grados de libertad para el brazo la Tabla 11, motor para el movimiento de brazo en la

Tabla 12, tecnología para el movimiento de cabeza en la Tabla 13, grados de libertad para el movimiento de cabeza en la Tabla 14, motor para el movimiento de cabeza en la Tabla 15, sistema de ojos en la Tabla 16, panel digital para los ojos en la Tabla 17 y fuente de alimentación en la Tabla 18.

Tabla 4
Módulo 1 (Material General)

Opción	Proceso de Manufactura	Peso	Precio	Orden de Selección
Plástico	Impresión 3D	Bajo	Bajo	1
Metálico	Fundición	Alto	Alto	2

Nota. Elaboración propia.

Tabla 5
Material para impresión 3D

Opción	Precio	Resistencia	Operabilidad	Orden de Selección
PLA	$25	Media	Fácil	1
ABS	$30	Alto	Difícil	2
TPU	$40	Baja	Fácil	3

Nota. Elaboración propia.

Tabla 6
Módulo 2 (Tarjeta de control principal)

Opción	Precio	RAM	ROM	Orden de Selección
Raspberry Pi	$180	16gb	128gb	1
Jetson nano	$230	4gb	16gb	2

Nota. Elaboración propia.

Tabla 7
Tarjeta de control secundaria

Opción	Precio	Conexión Inalámbrica	Tamaño	Peso	Oren de Selección
Esp32	$17	Si	68.6 mm x 53.4 mm	6g	1
Arduino UNO	$27	No	52 mm x 25 mm	25g	2

Nota. Elaboración propia.

Tabla 8
Módulo 3 (Visualización)

Opción	Precio	Interacción	Operabilidad	Orden de Selección
Pantalla Táctil	Alto	Alta	Fácil	1
Pantalla (No táctil)	Medio	Baja	Difícil	2

Nota. Elaboración propia.

Tabla 9
Modelo de pantalla

Opción	Precio	Tamaño	Altavoces	Compatibilidad	Orden de Selección
iPistBit	$70	10.1”	SI	Raspberry Pi 5/4/3/Zero	1
Gallagher	$130	10.1”	SI	Raspberry Pi 5/4B/3B+/3B/B+ Jetson Nano Win11/10	2
iUoyoMon	$40	7”	NO	Raspberry Pi 5/4/3/B+/A+	3

Nota. Elaboración propia.

Tabla 10
Módulo 5 (Movimiento de brazo): Tecnología

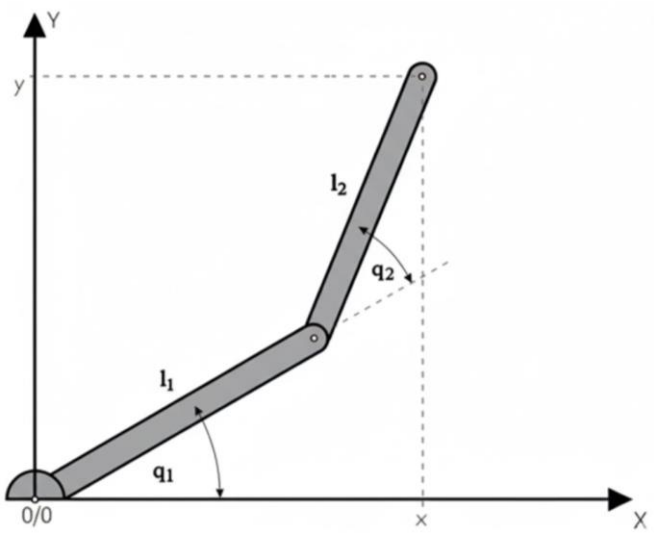
Opción	Fuerza	Control	Voltaje	Orden de Selección
Servomotor	Alta	Fácil	5V	1

Motor a pasos	Medio	Alto	12V	2
Motor DC	Baja	Fácil	3V	3

Nota. Elaboración propia.

El módulo 5 representa el movimiento básico de los brazos, el cual será de ayuda para que el sistema sea más dinámico. En base a esta necesidad se plantean las opciones de solución con 1, 3 o 2 grados de libertad como se muestra en la Figura 6.

Figura 6
Sistema de movimiento de 2 grados de libertad



Nota. Obtenido de (Chaparro Altamirano et al., 2015).

Tabla 11
Módulo 5 (Movimiento de brazo): Grados de libertad

Opción	Complejidad	Control	Facilidad de Montaje	Orden de Selección
1	Nula	Fácil	Alta	1
2	Medio	Mediano	Media	2
3	Alta	Extremo Difícil	Baja	3

Nota. Elaboración propia.

Tabla 12
Motor para el movimiento de brazo

Opción	Complejidad	Peso	Tamaño	Orden de Selección
Sg90	Nula	Poco	22 x 31 x 12,5 mm	1
Sg995	Nula	Mediano	58 x 36.5 x 20 mm	2

Nota. Elaboración propia.

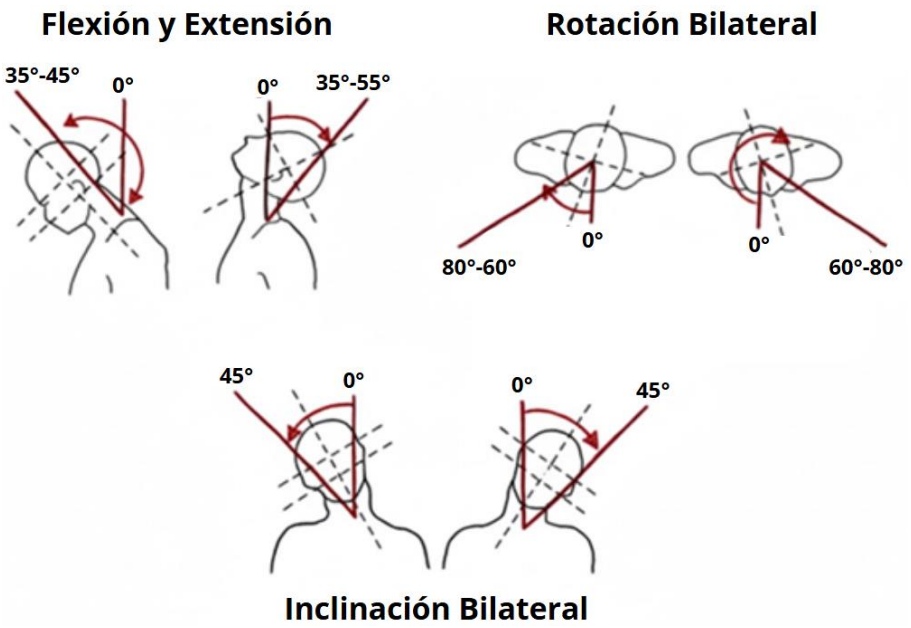
Tabla 13
Módulo 6 (Movimiento de cabeza): Tecnología

Opción	Fuerza	Control	Voltaje	Orden de Selección
Servomotor	Alta	Fácil	5V	1
Motor a pasos	Medio	Alto	12V	2
Motor DC	Baja	Fácil	3V	3

Nota. Elaboración propia.

Dado que el movimiento de la cabeza es un tema algo complejo se presentan en la Figura 7 las opciones de movimiento anatómicas existentes de las cuales se reducirá los grados de libertad necesarios en este movimiento para el robot.

Figura 7
Movimientos anatómicos de la cabeza.



Nota. Extraído de (Chiara, 2021).

Tabla 14
Módulo 6 (Movimiento de cabeza): Grados de libertad

Opción	Complejidad	Control	Facilidad de Montaje	Diseño estético	Orden de Selección
1	Media	Fácil	Media	Mediano	1
2	Alta	Dificil	Baja	Alto	2

Nota. Elaboración propia.

Tabla 15
Motor para el movimiento de cabeza

Opción	Complejidad	Peso	Tamaño	Orden de Selección
Sg995	Nula	Mediano	22 x 31 x 12,5 mm	1
Sg90	Nula	Poco	58 x 36.5 x 20 mm	2

Nota. Elaboración propia.

Tabla 16
Módulo 7 (Sistema de Ojos)

Opción	Complejidad	Control	Facilidad de Montaje	Diseño estético	Orden de Selección
Ojos Digitales	Baja	Fácil	Media	Alto	1
Ojos mecánicos	Alta	Difícil	Baja	Medio	2

Nota. Elaboración propia.

Tabla 17
Panel digital para los ojos

Opción	Complejidad	Control	Facilidad de Montaje	Personalización de gestos	Orden de Selección
Matriz de LED 8x8	Baja	Medio	Alta	Baja	1
Matriz de LED 8x32	Alta	Difícil	Media	Medio	2
Mirada Fija (Calcomanía)	Nula	Nulo	Alto	Nula	3

Nota. Elaboración propia.

Tabla 18
Módulo 7 (Fuente de alimentación)

Opción	Movilidad	Simplicidad	Potencia	Requerimientos adicionales	Orden de Selección
Adaptador DC	No	Alta	Alta	No	1
Batería Lipo	Si	Baja	Baja	Si	2
Baterías 26650	Si	Baja	Baja	Si	3

Nota. Elaboración propia.

Generación de Conceptos

Este apartado muestra las propuestas conceptuales elaboradas para determinar la configuración global del robot educativo, además de integrar criterios de tipo estético y funcional provenientes de los requerimientos previamente identificados. Cada concepto explica una combinación diferente de elementos mecánicos, electrónicos y de interacción, los cuales permiten ver distintas alternativas viables respecto a la estructura, a los actuadores, a la interfaz gráfica y al sistema de control. La descripción detallada de las características técnicas y del enfoque de diseño de cada propuesta establece, así, la base comparativa necesaria para la selección posterior del concepto más adecuado para el prototipo final.

Concepto 1

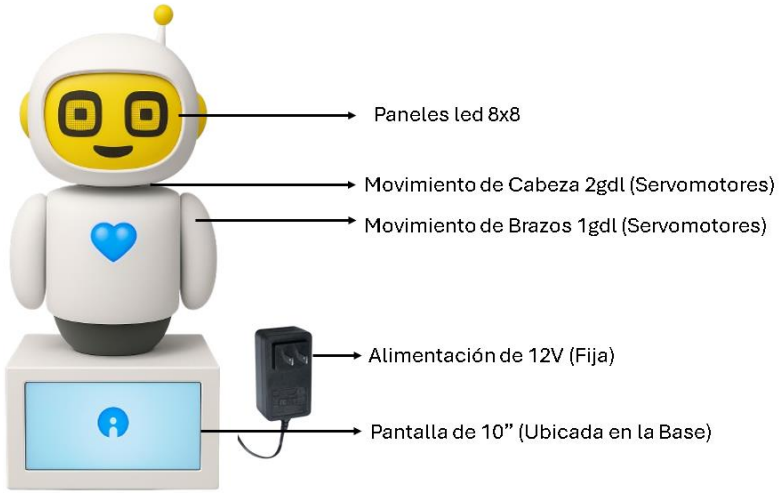
Robot educativo presentado en la Figura 8 de aspecto amigable y proporciones equilibradas, con una estructura impresa en PLA mediante impresión 3D, superficie lisa y colores suaves (blanco y celeste). Está diseñado para ofrecer una interacción visual y táctil destacada, priorizando la interfaz gráfica y el contacto directo con el usuario infantil.

Componentes técnicos principales:

- **Control principal:** Raspberry Pi (16 GB RAM / 128 GB ROM).
- **Control secundario:** ESP32, que se dedica al manejo de servomotores y LEDs.
- **Pantalla táctil:** iPistBit de 10.1” (interfaz HMI principal).
- **Ubicación de pantalla:** En la base adelantada de la inclinación de 30°, introducida en un soporte ergonómico para que el infante interactúe en un estímulo sentándose o de pie.
- **Movimiento de brazos:** Servomotores SG90 (1 grado de libertad) permiten un movimiento suave y acompañado de saludo o acompañamiento de voz.

- **Movimiento de cabeza:** Servomotor SG995, con 2 grados de libertad (inclinación y giro).
- **Ojos digitales:** Matriz LED 8×8, simulando expresiones básicas.
- **Audio:** Parlante y micrófonos integrados en el torso superior.
- **Fuente de energía:** Transformador AC (110V) a DC (12v), uso estacionario.
- **Tamaño:** 50 de alto x 25 de ancho

Figura 8
Referencia del concepto 1.



Nota. Elaboración con IA.

Concepto 2

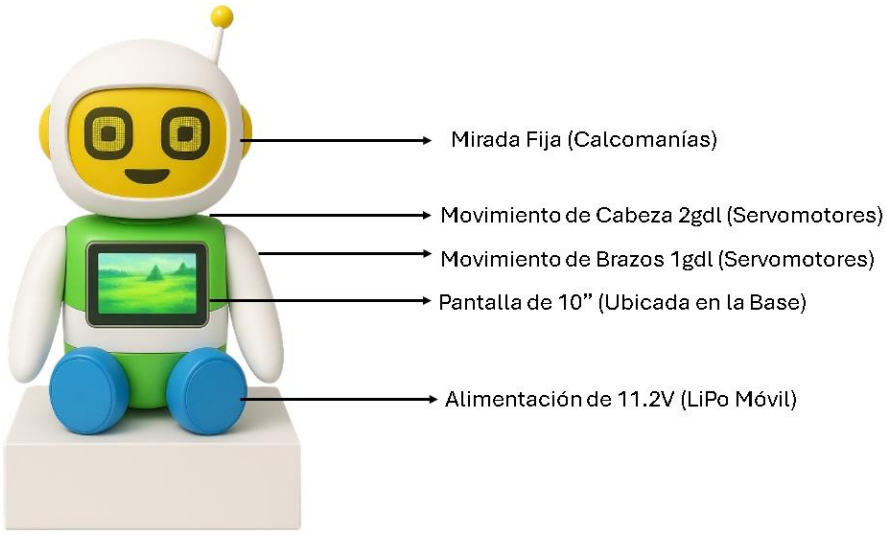
Robot mostrado en la Figura 9 construido de manera compacta y ligera, con una alta portabilidad. La pantalla de 7" se inserta en el torso, favoreciendo la interacción directa. Presenta un sistema redondeado en la que se visualiza una cabeza grande y unos brazos reducidos, que evocan ternura y proximidad entre ambos.

Componentes técnicos principales:

- **Material estructural:** PLA impreso en 3D.

- **Control principal:** Jetson Nano (4 GB RAM / 16 GB ROM).
- **Control secundario:** ESP32 incluye conexión inalámbrica y sensores sencillos.
- **Pantalla táctil:** iUoyoMon 10” sin altavoces; menor coste.
- **Ubicación de pantalla:** En el torso frontal, en posición vertical.
- **Movimiento de brazos:** Servomotores SG995, 2 grados de libertad, para llevar a cabo gestos amplios (abrir, señalar o saludar).
- **Movimiento de cabeza:** Servomotor SG90, 1 grado de libertad (inclinación frontal simple).
- **Ojos digitales:** Calcomanía fija de retroiluminación (opción 3 de módulo 7)
- **Fuente de energía:** Batería LiPo recargable (con la meta de movilidad en aula).

Figura 9
Referencia del concepto 2.



Nota. Elaborado con IA.

Con los resultados de la Tabla 20 con respecto a las valorizaciones de la Tabla 19 se escoge el primer concepto ya que destaca de manera específica en costo y durabilidad dado por la posición de la pantalla en la parte baja.

Tabla 19
Valorización de conceptos

Criterio	1 a 3	4 a 7	8 a 9
Visualización	Mala	Buena	Excelente
Procesamiento	Lento	Medio	Rápido
Durabilidad	Frágil	Medianamente Robusto	Robusto
Transportabilidad	Mala	Normal	Excelente
Construcción	Difícil	Normal	Fácil
Estética de diseño	Diseño	Diseño Básico	Diseño Amigable
Costo	Costosa	Normal	Barato

Nota. Elaboración propia.

Tabla 20
Selección de conceptos

Criterio	Valor	Concepto 1	Concepto 2
Visualización	20	8	9
Procesamiento	10	3	7
Durabilidad	20	9	3
Transportabilidad	5	7	5
Construcción	5	7	8
Estética de diseño	20	9	8
Costo	20	9	3
Total	**100**	**52**	**43**

Nota. Elaboración propia.

Diseño del Prototipo

El diseño del prototipo se ha dividido en los sistemas principales del mismo, los cuales consisten en el sistema mecánica, donde se especifican todos los elementos físicos a utilizar y los principales sistemas para el movimiento de la cabeza, y los brazos, el sistema eléctrico y de control, que especifica las conexiones eléctricas que existentes entre los distintos componentes electrónicos existentes, así como el sistema que controla todos estos elementos, finalizando con el diseño del software general y el HMI que define la interactividad entre el usuario y el robot, cumpliendo con los requisitos funcionales solicitados.

Sistema Mecánico

A continuación, se explican las soluciones mecánicas del robot pedagógico, enfatizando a continuación cómo las partes físicas sirven para sostener y guiar la pieza, y la manera de transmitir el movimiento para poder enlazarlas posteriormente con las actividades a realizar con los niños de 4-5 años. El diseño propuesto se pretende simple, resistente y fácil de montar y desmontar: dos brazos, un grado de libertad cada uno, una cabeza y dos grados de libertad, para terminar con una base rígida que, además, engloba, la pantalla y la electrónica correspondiente.

Movimiento de brazos

El movimiento de los brazos se hace mediante servos, los cuales están directamente unidos al brazo mediante un accesorio (el brazo del servo) y un tornillo central que asegura el acoplamiento entre el servo y el eje. El servomotor queda sostenido por un soporte de dos orejas que se atornilla a la estructura principal, de modo que la fuerza del motor sale del eje del servo, viaja a través del brazo y no encuentra otra conexión o componente intermedio al llegar a él. El soporte sostiene el peso de este conjunto y evita la rotación o el desplazamiento del servo en la base. El eje de giro del brazo se diseñó como un cuadro redondeado en vez de un eje totalmente cilíndrico para facilitar la manufactura por impresión 3D, ya que las piezas circulares requieren

soportes de impresión mientras que las superficies planas del modelo reducen el material de soporte y mejoran la firmeza de la pieza.

El sistema este hecho de manera simétrica lo que quiere decir que su equivalente en el lado izquierdo se realiza con las mimas piezas, simplemente ubicándolas en su nueva posición, reduciendo la complejidad del diseño final, tal como se observa en la Figura 10.

Figura 10
Sistema de movimiento de brazo.

Nota. Elaboración Propia.

Haciendo uso de las ecuaciones dinámicas procedemos a encontrar la inercia del eslabón.

$$I = m \cdot l^2 \quad (1)$$

Donde $m = 0{,}042\ kg$ y $l = 0{,}07m$:

$$I = 2{,}06 \times 10^{-4} kg \cdot m^2 \quad (2)$$

Aceleración y velocidad angular

$$\alpha_{max} = 4 \cdot \frac{\Delta\theta}{\Delta t^2} \quad (3)$$

$$\omega_{max} = 2 \cdot \frac{\Delta\theta}{\Delta t} \quad (4)$$

$$\alpha_max \approx 145\, rad/s^2 \quad (5)$$

$$\omega_max \approx 12{,}3\, rad/s \quad (6)$$

Torque inercial

$$\tau_{in} = I \cdot \alpha_{max} \quad (7)$$

Donde $I = 2{,}06 * 10\text{^} - 4\, kg * m^2$ y $\alpha_{max} = 145\, rad/s^2$

$$\tau_{in} \approx 0{,}0298\, N \cdot m \quad (8)$$

Torque por Gravedad

$$\tau_{g(\theta)} = m \cdot g \cdot l \cdot sen(\theta) \quad (9)$$

Donde $m = 0{,}042\, kg$, $l = 0{,}07\, m$ y $g = 9{,}81\, m/s^2$

$$\tau_{g(\theta)} \approx 0{,}0288\, N \cdot m \quad (10)$$

Torque Total

$$\tau_{total} = \tau_{in} + \tau_g \quad (11)$$

$$\tau_{total} \approx 0{,}6\, kg \cdot cm \quad (12)$$

Al tratarse de un elemento interactivo que va a estar en contacto con niños, se escogerá un factor de seguridad significativo de al menos 7, de modo que el torque final resultaría

$$\tau_{total} \approx 4.2\ kg \cdot cm \quad (13)$$

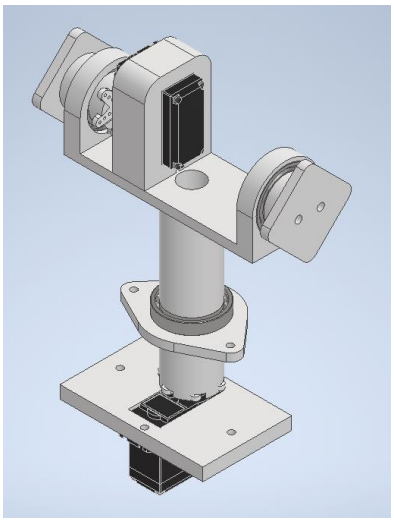
Lo que permite seleccionar los servomotores MG996R por su torque nominal de 9,4 kg*cm a 4,8V (hasta 11 kg*cm a 6V) y su precisión angular de 0,17 s/60°, lo que los hace adecuados para el tamaño y peso del prototipo.

Movimiento de cabeza

El sistema de movimiento de la cabeza se apoya en una columna central que nace en la placa base. En la unión entre la columna y la base se monta un rodamiento alojado en una brida, este rodamiento recibe todo el peso del conjunto superior y las fuerzas laterales cuando la cabeza gira. El servo de giro horizontal (yaw) queda fijado dentro de la base y solo aporta fuerza para la rotación.

En parte superior de la columna hay un diseño transversal que soporta la cabeza. Cada extremo de ese elemento se contiene un alojamiento para los rodamientos donde se instala el segundo movimiento, el paso o inclinación (pitch). El servo de pitch va atornillado al centro de la viga y transmite el giro a uno de los laterales mediante su *horn*. El otro lateral lleva un rodamiento espejo para cerrar el apoyo, de modo que la carga de la cabeza se reparte entre los dos lados y el servo, de nuevo, solo entrega torque sin soportar esfuerzos transversales, como se observa en la Figura 11.

Figura 11
Sistema de movimiento de cabeza.



Nota. Elaboración propia.

Bajo el mismo análisis, obtenemos los valores de torque correspondiente a los motores para este movimiento.

$$I = m \cdot l^2 \quad (14)$$

Donde $m = 0{,}2\ kg$ y $l = 0{,}025m$:

$$I \approx 1{,}25 \times 10^{-4} kg \cdot m^2 \quad (15)$$

Aceleración y velocidad angular

$$\alpha_{max} = 4 \cdot \frac{\Delta\theta}{\Delta t^2} \quad (16)$$

$$\omega_{max} = 2 \cdot \frac{\Delta\theta}{\Delta t} \quad (17)$$

$$\alpha_max \approx 145\ rad/s^2 \quad (18)$$

$$\omega_max \approx 12{,}3\ rad/s \quad (19)$$

Torque inercial

$$\tau_{in} = I \cdot \alpha_{max} \quad (20)$$

Donde $I = 1{,}25 * 10^{-4}\ kg * m^2$ y $\alpha_{max} = 145\ rad/s^2$

$$\tau_{in} \approx 0{,}0181\ N \cdot m \quad (21)$$

Torque por Gravedad

$$\tau_{g(\theta)} = m \cdot g \cdot l \cdot sen(\theta) \quad (22)$$

Donde $m = 0{,}2\ kg$, $l = 0{,}025\ m$ y $g = 9{,}81\ m/s^2$

$$\tau_{g(\theta)} \approx 0{,}0491\ N \cdot m \quad (23)$$

Torque Total

$$\tau_{total} = \tau_{in} + \tau_g \quad (24)$$

$$\tau_{total} \approx 0{,}7\ kg \cdot cm \quad (25)$$

Manteniendo el margen de seguridad de 7, se obtiene el torque final resultaría

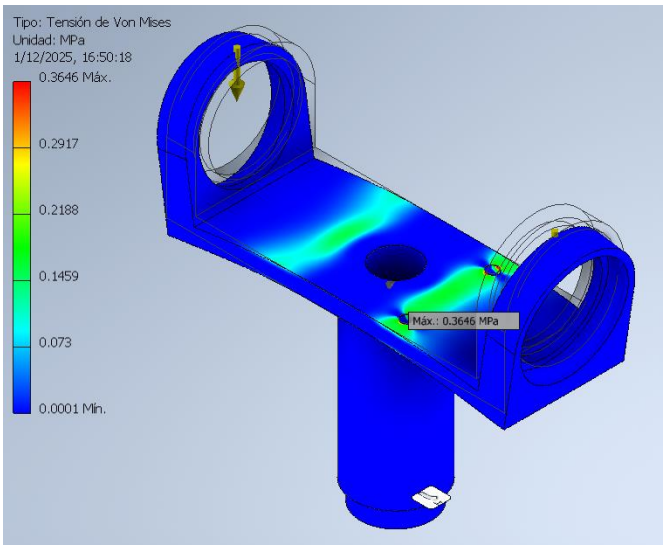
$$\tau_{total} \approx 4.9\ kg \cdot cm \quad (26)$$

Lo que permite mantiene la selección de los servomotores MG996R.

Dado que el eje de movimiento de la cabeza soporta todo el peso de la misma, es conveniente realizar un análisis de esfuerzos del sistema para corroborar su funcionalidad, como resultado obtenemos lo mostrado en la figura 12.

Figura 12

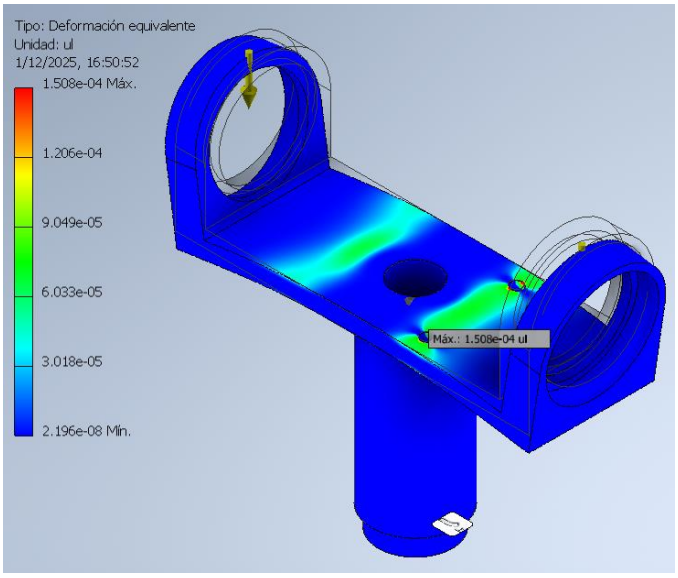
Tensión máxima que soporta el eje de la cabeza.



Nota. Elaboración propia.

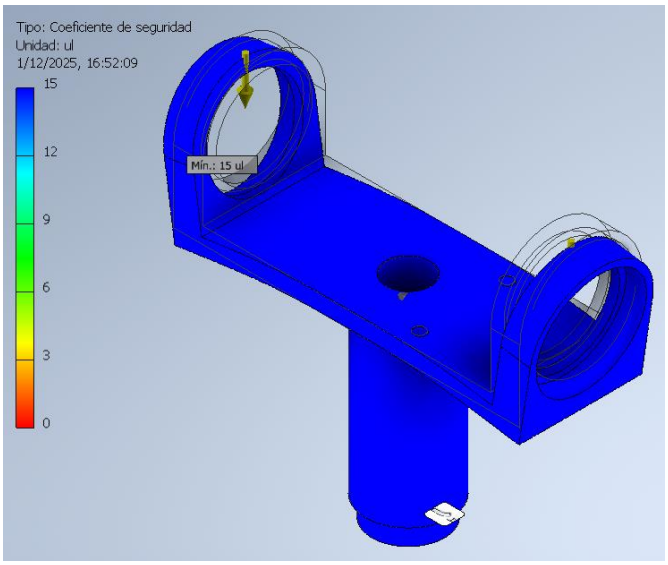
Figura 13

Deformación máxima ejercida en el eje de la cabeza.



Nota. Elaboración propia.

Figura 14
Coeficiente de Seguridad mínimo en el eje de la cabeza.



Nota. Elaboración propia.

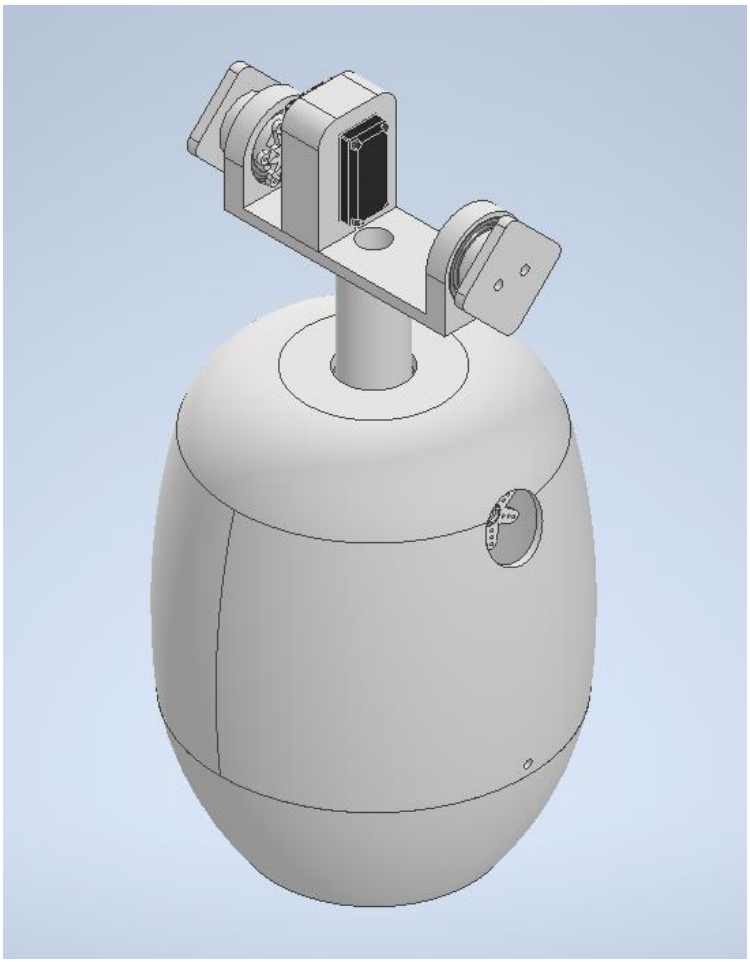
Obteniendo así un valor de esfuerzo máximo bajo las condiciones actuales de 0,366MPa que no supera el límite del material seleccionado.

Torso

El torso se diseñó en dos piezas: una inferior, que fija el cuerpo a la base, y una superior, que sostiene los conjuntos de giro de la cabeza y los brazos. Ambas piezas se unen por atornillado directo y dejan acceso libre a todos los puntos de sujeción para montaje manual de los sistemas mencionados. La pieza inferior actúa como anclaje estructural y reparte las cargas hacia la base; la superior integra los apoyos de los servos y rodamientos como se puede observar en la Figura 15.

Figura 15

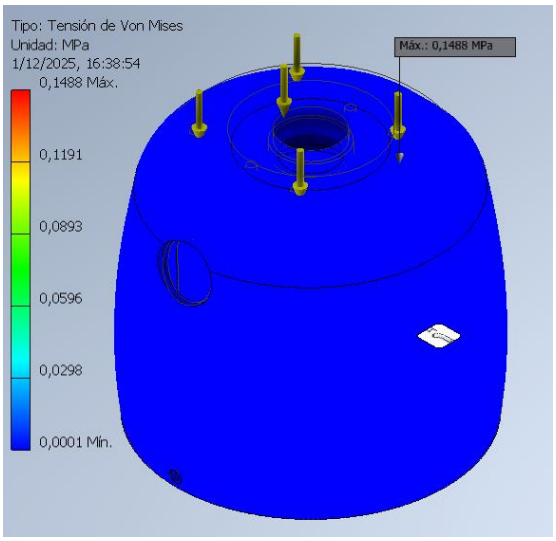
Torso que soporta los esfuerzos principales



Nota. Elaboración propia.

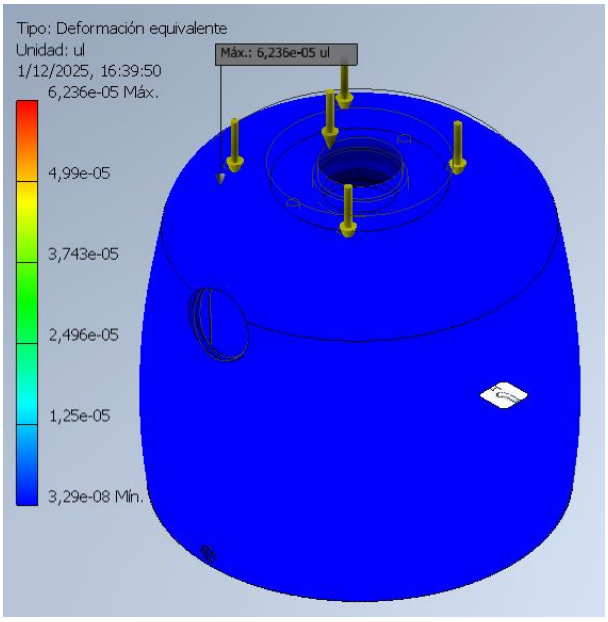
Realizando el análisis de esfuerzo correspondiente obtenemos los siguientes resultados:

Figura 16
Tensión máxima soportada por el Torso



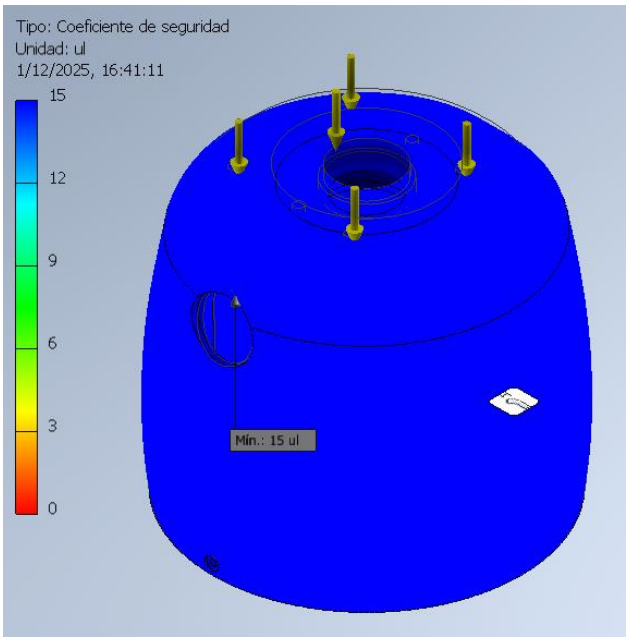
Nota. Elaboración propia.

Figura 17
Deformación máxima en el Torso



Nota. Elaboración propia.

Figura 18
Factor de Seguridad mínimo en el Torso



Nota. Elaboración propia.

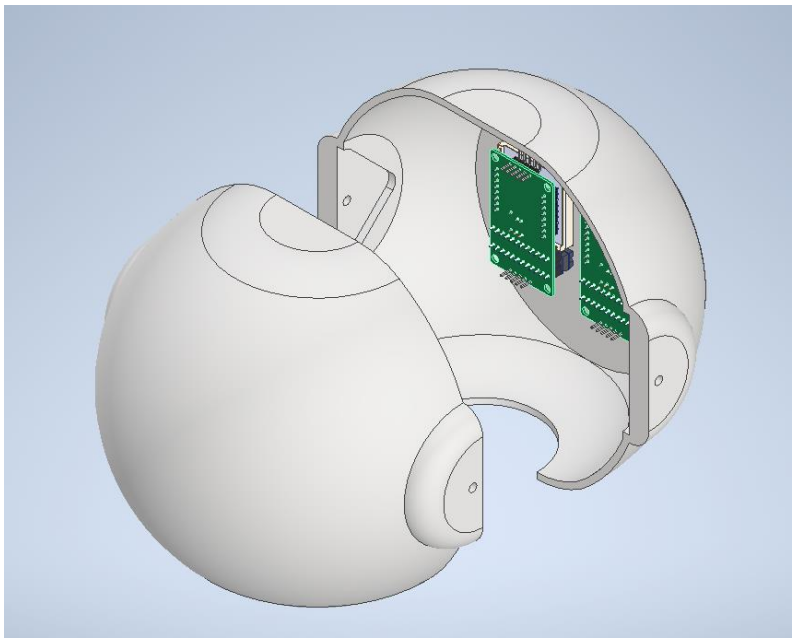
Como se muestra en la Figura 16, el esfuerzo máximo alcanzado es de 0,138 MPa, comparado con los 47 MPa, soportados por el PLA en tracción, hacen que este material sea el ideal manteniendo los costos bajos en el diseño, lo que se corrobora en el coeficiente de seguridad de 15 mostrado en la Figura 18.

Cabeza

La cabeza se dividió en dos partes, como se puede observar en la Figura 19 como una forma de garantizar el correcto ensamblaje de la misma, pues la estructura interna para el giro requiere su espacio correspondiente, además en la parte plana frontal que tiene el rol de rostro, se diseñaron alojamientos para las matrices led de 8x8 que lograrán integrar los gestos del robot para mantener su interactividad.

Se ajusta con dos pernos distribuidos en cada pieza que conforma la cabeza al eje central del movimiento, de esta manera garantizamos que no exista ningún juego en la unión y se garantice el movimiento de los componentes.

Figura 19
Cabeza con visualización de sus matrices led



Nota. Elaboración propia.

Visualización general

En conjunto todas las piezas se unen para formar la estructura principal del robot mostrado en la Figura 20, la cual es soportada en una base donde se ubicarán todos los gadgets, elementos de control relacionados con el movimiento de control y la pantalla principal, esto con el fin de reducir el estrés generado por estos componentes de manera directa en el robot y agregar nuevos periféricos al mismo.

Figura 20
Visualización general del robot

Nota. Elaboración propia.

Sistema Electrónico y de Control

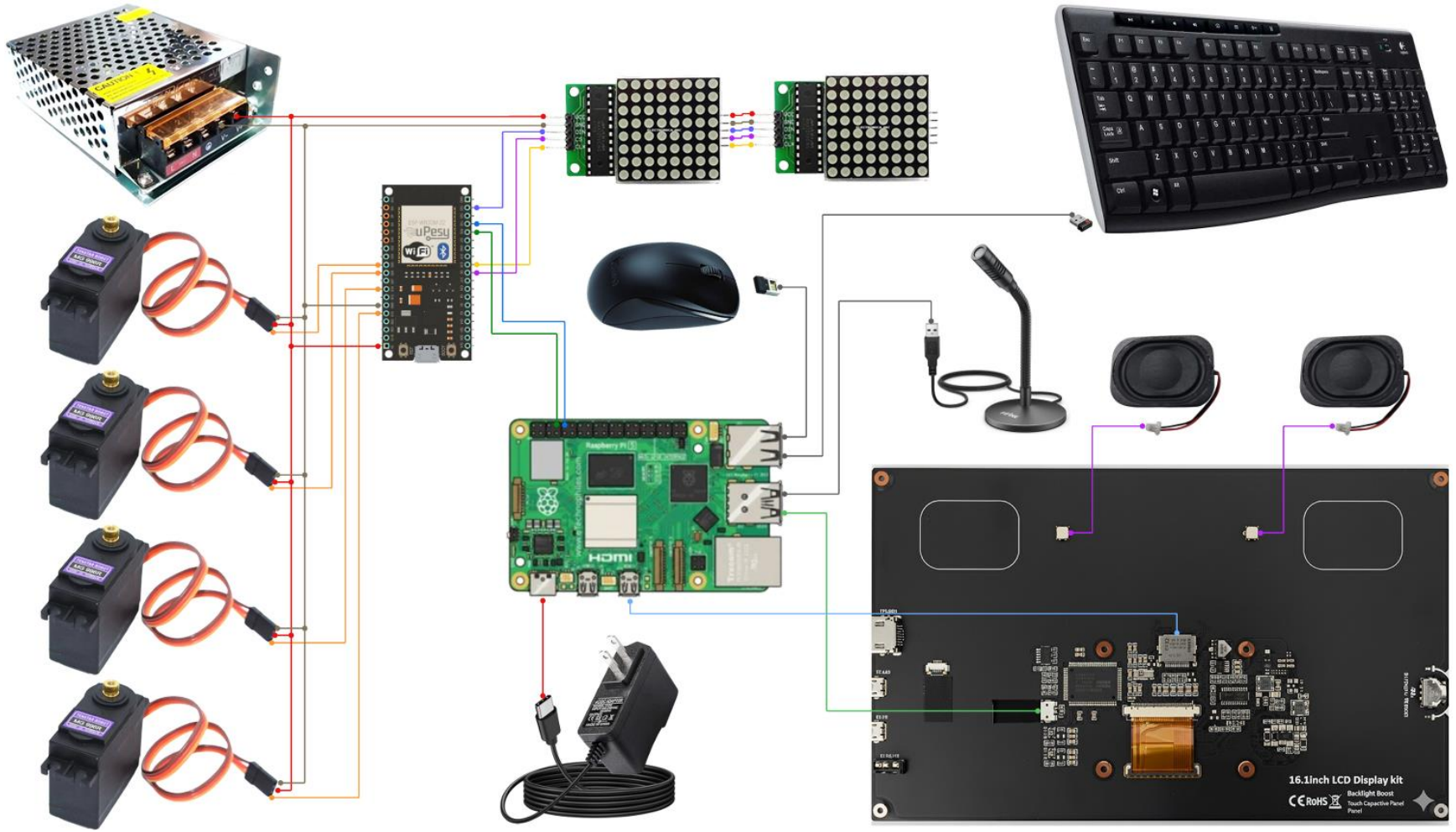
Se da cuenta de los aspectos correspondientes al diseño tanto electrónico como de control del robot educativo en esta sección, ofreciendo las ideas sobre la arquitectura general del sistema, la distribución de sus unidades de procesamiento, sus circuitos de conexión y de alimentación, la integración de sensores, actuadores y periféricos. También se ofrecen las configuraciones que se han utilizado para el correcto funcionamiento de la estabilidad eléctrica, la comunicación sincrónica entre módulos y la seguridad operativa del sistema, y con ello garantizar un funcionamiento correcto del prototipo.

Arquitectura Electrónica General

El sistema electrónico del robot educativo se diseñó con una arquitectura modular e integrada, orientada a garantizar la interactividad en tiempo real entre el niño y el prototipo. La estructura general la conforman dos unidades de control principales, una Raspberry Pi 5 y una ESP32, que trabajan de manera coordinada para integrar tanto las funciones cognitivas y de interfaz como el contr ol físico del robot.

En la Figura 21 se puede observar el diagrama de conexiones general para el robot:

Figura 21
Diagrama General de Conexiones



Nota. Elaboración propia.

Unidades de control

La arquitectura del sistema tiene una estructura modular de doble unidad de control para diferenciar el procesamiento cognitivo de la gestión del hardware crítico de baja latencia.

Unidad Central (Raspberry Pi 5)

- **Función principal:** Realiza las funciones de unidad central de procesamiento, es decir, controla el motor central de interacción lingüística.
- **Capacidades:** Esta unidad ha sido seleccionada por su alta capacidad de cómputo y por la posibilidad de interaccionar con librerías de reconocimiento de voz, análisis fonético y síntesis de texto a voz.
- **Elementos que ejecuta/controla:** La interfaz gráfica de usuario (GUI) en una pantalla táctil de 10.1 pulgadas, la tecnología TTS, STT y periféricos de entrada/salida como:
    - Micrófono USB
    - Bocinas
- **Conectividad adicional:** Salida micro HDMI para proyección externa.
- **Alimentación:** Puerto USB-C con adaptador regulado de 5V.

Unidad de Control Auxiliar (ESP32)

- **Función principal:** Actúa como unidad de control embebida para los actuadores y periféricos de baja latencia del robot (control físico).
- **Comunicación:** Conecta a la Raspberry Pi 5, mediante la creación de un enlace serial UART con un voltaje de 3.3 V, garantizando una baja latencia y una sincronización de estímulos.

- **Elementos que controla directamente:**
    - 4 servomotores MG996R.
    - 2 matrices LED 8x8 (en cascada).

Sensores

Un grupo de sensores recoge la interacción del usuario y proporciona información de retorno en tiempo real mientras se están realizando las actividades de lenguaje.

Micrófono USB

Se conecta de manera directa a la Raspberry Pi 5 y captura la voz del niño, operando en los escenarios del análisis fonético (detección de la articulación correcta de las palabras) y la comprensión oral.

Pantalla Táctil (10.1")

Se conecta mediante Micro HDMI (video) y USB (táctil)

- **Entrada:** Superficie multitáctil para interacción directa (selección, navegación).
- **Salida:** Visualización de la GUI y contenido gráfico (1024×600 px).

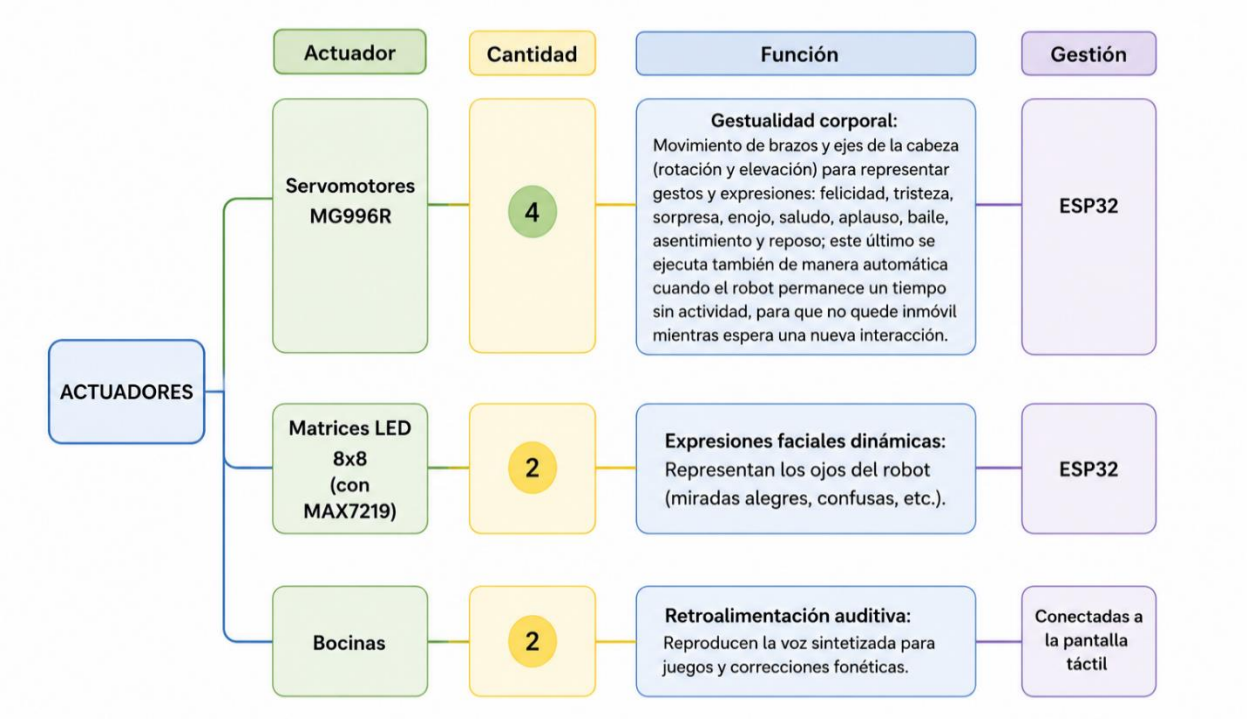
Bocinas Integradas

Se conectan mediante cable JST a la pantalla, proporcionando retroalimentación auditiva directa (voz sintetizada y sonidos) desde la unidad de visualización.

Actuadores

El sistema utiliza actuadores gestionados por la ESP32 para proporcionar expresividad física y retroalimentación visual y auditiva, lo cual se puede observar en la **¡Error! No se encuentra el origen de la referencia.**.

Figura 22
Actuadores utilizados

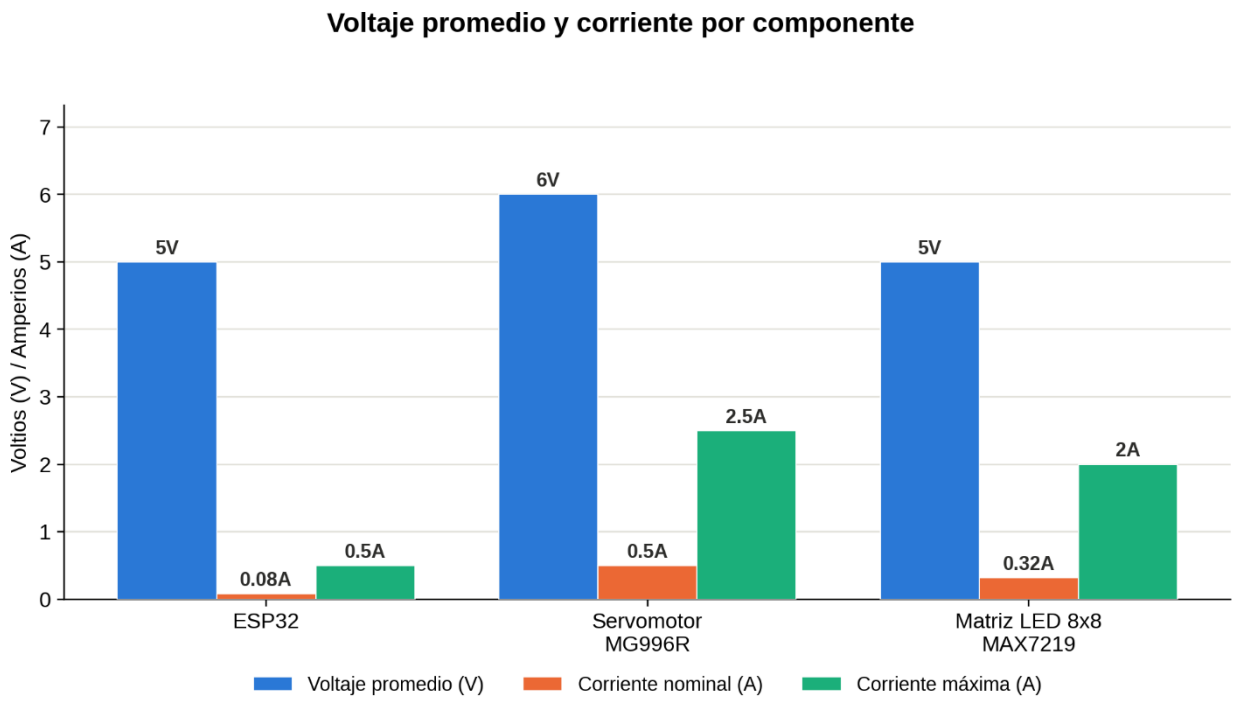


Nota. Elaboración propia.

Sistema de Alimentación

Para poder seleccionar un sistema de alimentación primero se consultarán las corrientes y voltajes a los que trabajan los diferentes elementos que conforman el robot, como se muestra en la **¡Error! No se encuentra el origen de la referencia.**.

Figura 23
Consumo de voltaje y corriente



Nota. Elaboración propia.

Se necesita una alimentación capaz de proporcionar hasta 500 mA según sus especificaciones que la ESP32 requiere para cubrir los picos de consumo generados durante el uso de Bluetooth y Wi-Fi, aunque su consumo nominal se mantiene alrededor de 80 mA. Los servomotores, por su parte, presentan una corriente nominal cercana a 500 mA durante el movimiento, pero pueden demandar hasta 2.5 A en condición de bloqueo. Las matrices LED, finalmente, registran un consumo nominal de aproximadamente 320 mA, aunque con brillo máximo y todos los LED activados la corriente puede subir hasta 2 A.

Con estas consideraciones, al calcular las corrientes nominales multiplicando por la cantidad de cada elemento se obtiene que:

$$I_{Total} = I_A \cdot N_A + I_B \cdot N_B + I_C \cdot N_C \quad (27)$$

En donde:

- $I_A = 80mA$
- $N_A = 1$
- $I_B = 500mA$
- $N_B = 4$
- $I_C = 320mA$
- $N_C = 2$

Se tiene entonces que la corriente total con los valores nominales es:

$$I_{Total_nominal} = 2720\ [mA] = 2.72\ [A] \quad (28)$$

Sin embargo, también realizaremos el cálculo para las corrientes máximas para poder trabajar con ellas en caso de picos de corriente inesperados y poder tener un factor de seguridad alto:

$$I_{Total} = I_A \cdot N_A + I_B \cdot N_B + I_C \cdot N_C \quad (29)$$

En donde:

- $I_A = 500mA$
- $N_A = 1$
- $I_B = 2500mA$
- $N_B = 4$
- $I_C = 2000mA$
- $N_C = 2$

Se tiene entonces que la corriente total con los valores máximos es:

$$I_{Total} = 14500\ [mA] = 14.5\ [A] \quad (30)$$

Tomando este valor como referencia, se determina que es necesario emplear una fuente capaz de suministrar 5V y 15A, con el fin de evitar caídas de tensión o limitaciones de corriente en situaciones donde los componentes operen a su máxima demanda. De este modo, calcularemos los factores de seguridad.

Factor de seguridad nominal:

$$FS_{nominal} = \frac{Fuente}{Consumo\ Nominal} \quad (31)$$

En donde:

- $Fuente = 15A$
- $Consumo\ Nominal = 2.72A$

$$FS_{nominal} = 5.51 \quad (32)$$

Factor de seguridad máximo:

En donde:

- $Fuente = 15A$
- $Consumo\ Nominal = 2.72A$

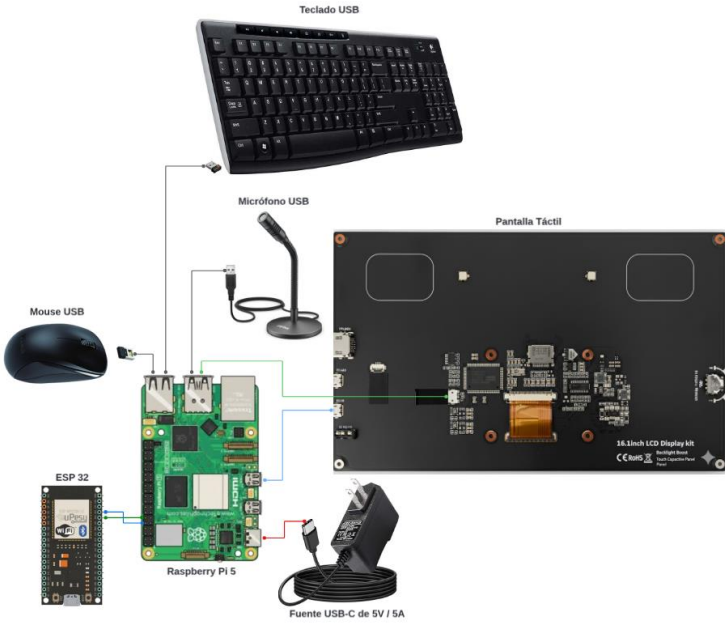
$$FS_{máx} = 1.03 \quad (33)$$

Se obtiene un factor de seguridad de 5.51 con respecto al consumo nominal de 2.72 A, así como un factor de seguridad de 1.03 frente al consumo máximo estimado de 14.5 A. En consecuencia, queda un margen aproximado de 500 mA incluso en un escenario donde todos los componentes operen de manera simultánea y a su máxima capacidad, lo que permite prevenir posibles fallos por insuficiencia de corriente.

Sistema Lógico (Raspberry Pi 5)

La Figura 24 muestra el diagrama del sistema del robot educativo, con la distribución de energía desde una fuente hacia la Raspberry, la pantalla táctil, los periféricos y la comunicación con la ESP32.

Figura 24
Conexiones del sistema Lógico

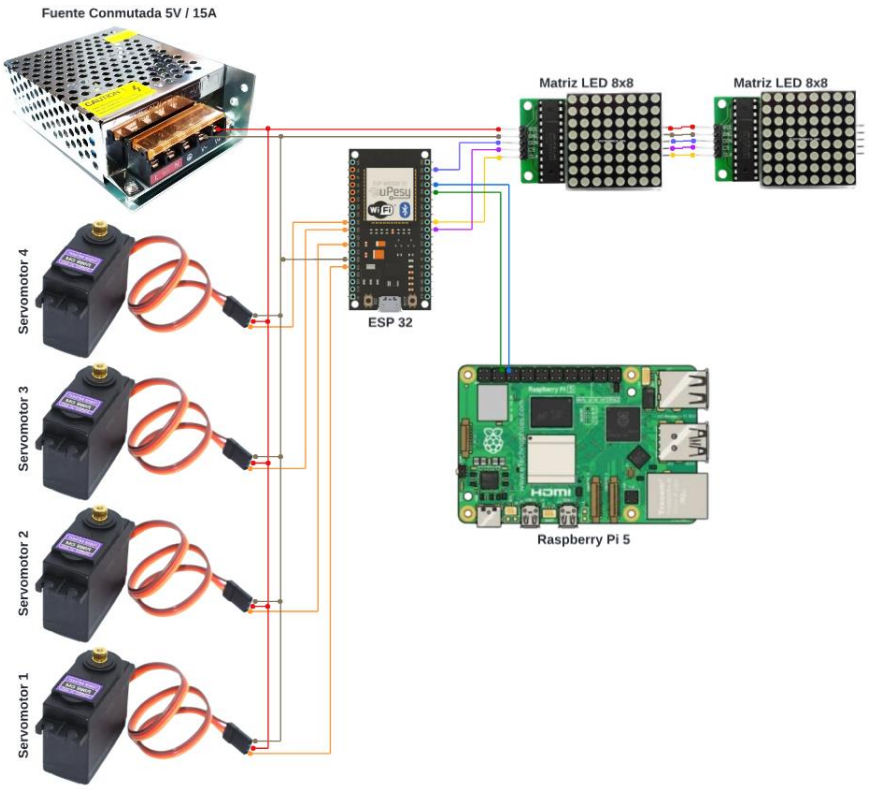


Nota. Elaboración propia.

Sistema de Potencia (Actuadores y ESP32)

La Figura 25 muestra el diagrama lógico del sistema de potencia del robot educativo. La energía es suministrada por una fuente conmutada, la cual alimenta a la ESP32, los cuatro servomotores y las dos matrices LED.

Figura 25
Conexiones del Sistema de Potencia



Nota. Elaboración propia.

Circuitos de control y conexión

La estructura de control se ha pensado para un diseño modular y directo con el objetivo de dotar a la comunicación de estabilidad, facilidad de montaje y seguridad desde el punto de vista eléctrico.

Estructura de Control

1. **Raspberry Pi 5 (RPi 5):** Unidad de procesamiento (voz y GUI).
2. **ESP32:** Unidad de control de actuadores físicos y módulos visuales.

Comunicación Serial

Ambos procesadores se comunican mediante un cable USB, por donde la Raspberry Pi envía órdenes a la ESP32, utilizando comunicación serial. Al iniciar el sistema, el software identifica automáticamente en qué puerto está conectada la ESP32 —reconoce los adaptadores USB-serial más comunes en este tipo de placas— en lugar de requerir que el puerto se configure manualmente cada vez. Si llegara a perderse el enlace, el sistema trata de conectarse otra vez, de manera automática, sin tener que reiniciar la aplicación.

- **Interconexión:** Se establece por medio del cable USB entre la Raspberry Pi 5 y la ESP32 mediante una velocidad de transmisión de 9600 baudios.

Control de Actuadores (Gestionados por ESP32)

- **Servomotores MG996R (4):** Cada uno es controlado individualmente mediante señales PWM Ea/g con GPIO (25, 26, 14, 13).
- **Conjunto de matrices LED 8x8 (2):** En una conexión en cascada a pines GPIO específicos (DIN en 23, CS en 5, CLK en 18) de módulos MAX7219.

Esquema de Alimentación

- **Sistema de Potencia (ESP32 y Actuadores):** Fuente independiente de 5V/15A.
- **Sistema Lógico (RPi 5 y Periféricos):** Fuente independiente de 5V/5A por USB-C.
- **Seguridad:** Los dos circuitos de alimentación comparten una tierra común y el sistema de picos de corriente utiliza condensadores de desacoplo.

Conexión de Periféricos

- Los elementos de entrada/salida (micrófono, pantalla y táctil) se pueden conectar a los puertos USB de la Raspberry Pi 5 directamente.

Diseño PCB

El diseño se centró en la sección de potencia del circuito, abarcando las conexiones necesarias para la ESP32 y todos los actuadores que esta controla, junto con sus respectivas alimentaciones. Se integró un regulador de voltaje con el fin de suministrar la alimentación requerida a la ESP32, que opera a 3.3V.

Para el cálculo del grosor de las pistas conductoras, se tiene que:

$$W = \left(\frac{I}{K \cdot (\Delta T)^B \cdot t}\right)^{\frac{1}{C}} \quad (34)$$

En donde:

- Aumento de temperatura: ΔT = 15°C
- Espesor de cobre: t = 1 oz
- Constantes para pista externa:
  - K = 0.048
  - B = 0.44
  - C = 0.725

Ahora se calculará para cada corriente máxima:

- Para $I = 500mA,$ en la etapa de alimentación para la ESP32.

$$W = 4.90mils \quad (35)$$

- Para $I = 14A,$ en la etapa de alimentación para los 4 Servomotores MG 996R y las 2 matrices LED.

$$W = 485.38mils \quad (36)$$

- Para $I = 40mA,$ en las señales de control de la ESP32.

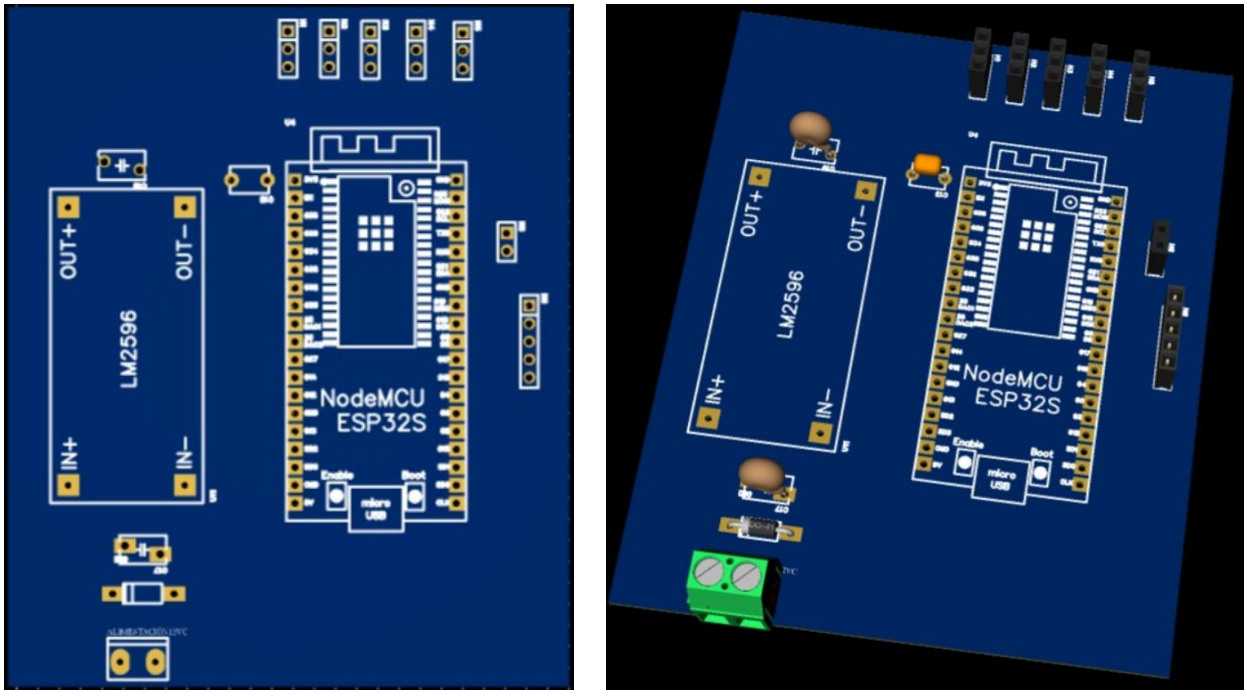
$$W = 0.15mils \tag{37}$$

- Para $I = 14.5A,$ en la etapa de la Fuente de alimentación a los componentes.

$$W = 509.45mils \tag{38}$$

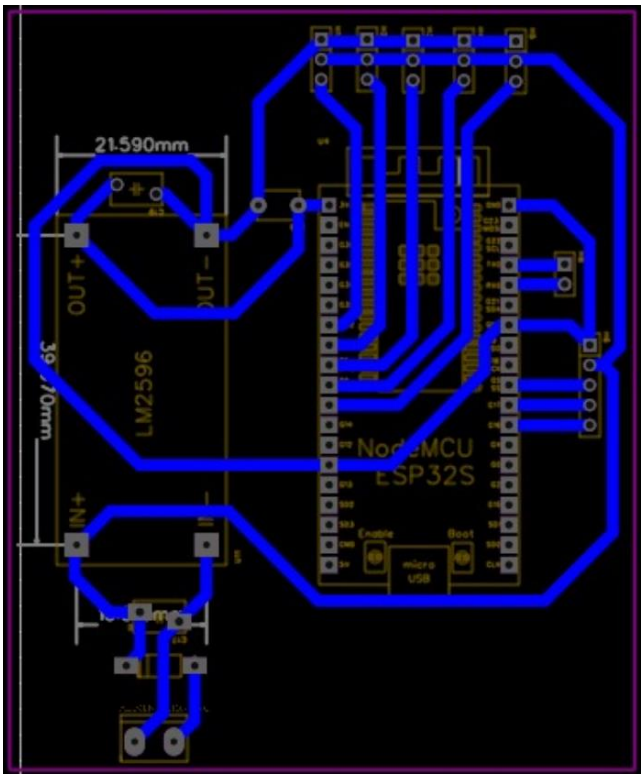
El diseño de la PCB se puede observar en las Figura 26, Figura 27 y Figura 28

Figura 26
Diseño PCB para el proyecto.



Nota. Elaboración propia.

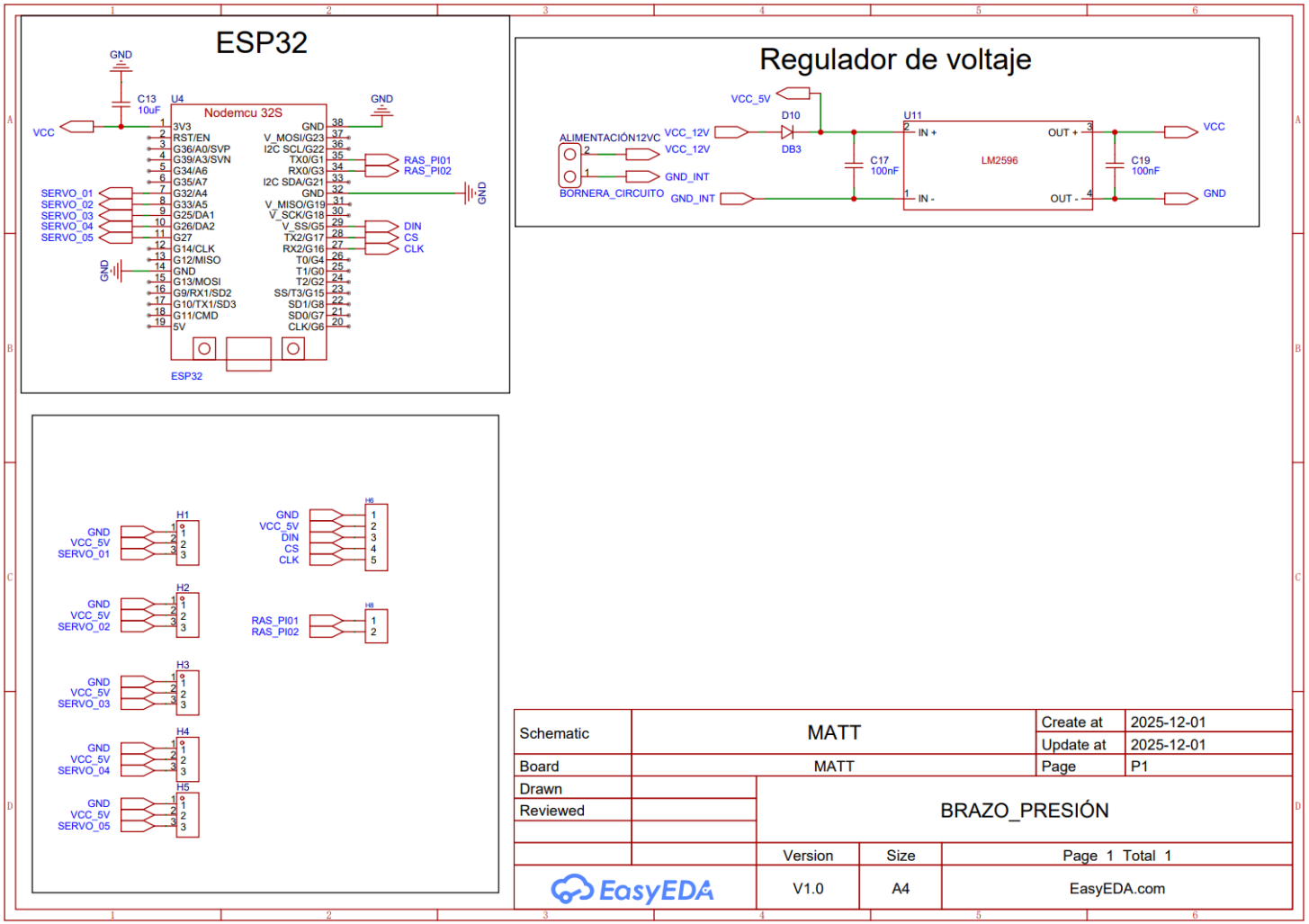
Figura 27
Diseño de pistas PCB.



Nota. Elaboración propia.

Figura 28

Conexiones para diseño de PCB.



Nota. Elaboración propia.

Interfaz Gráfica

La arquitectura de la interfaz gráfica del prototipo robótico se diseñó con un enfoque modular orientado a la interacción infantil, con las limitaciones físicas del hardware embebido y los requerimientos pedagógicos del proyecto como referencia. La GUI se implementó en Python: Qt Designer para el diseño de ventanas y PyQt5 como framework para la lógica y el renderizado de la interfaz. Esta arquitectura mantiene una separación clara entre la capa visual y la lógica de control, lo que favorece la mantenibilidad, la escalabilidad y la reutilización del código.

Navegación

En la Figura 29, se puede observar la distribución general para la navegación que se usa para los distintos juegos para esta interfaz gráfica:

Figura 29
Diagrama General de la navegación de la HMI.

Nota. Elaboración propia.

Arquitectura

En esta sección se describen de manera detallada las distintas pantallas del sistema, especificando su estructura, el contenido que presentan y la función de cada uno de sus elementos, así como las transiciones y flujos de navegación asociados a cada interacción.

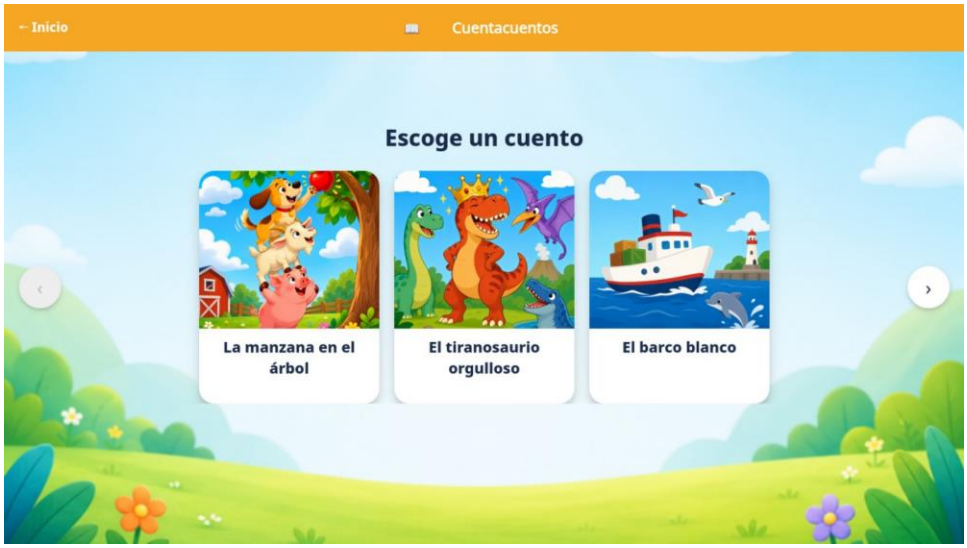
En la Figura 30 se muestra la pantalla correspondiente al menú principal del sistema. En la parte superior se dispone un elemento tipo label pregunta al usuario la acción de seleccionar una actividad. Posteriormente, se presentan tres pushbuttons asociados a los módulos interactivos disponibles: Cuenta Cuentos, Ordena Oraciones y Pronunciación, cada uno de los cuales redirige a la ventana específica de su respectivo juego. Adicionalmente, se incluyen dos pushbutton, el primero con iconografía de equis (X) el cual ayudará a salir de la aplicación y otro con un símbolo de engranaje para poder acceder a las opciones y configuraciones.

Figura 30
Pantalla de menú de juegos de la HMI

Nota. Elaboración propia.

En la Figura 31 se muestra la ventana correspondiente al menú de cuentos, la cual se despliega al seleccionar el módulo Cuenta Cuentos. En esta interfaz se incorpora un *label* instructivo que indica al usuario la selección de uno de los doce cuentos disponibles. Debajo de dicho elemento se dispone un menú de selección conformado por doce *pushbuttons*, cada uno asociado a un cuento específico. Finalmente, en la parte inferior de la ventana se incluye un *pushbutton* con iconografía de flecha de retroceso, cuya función es retornar al menú de selección de juegos.

Figura 31
Pantalla de menú de cuentos de la HMI

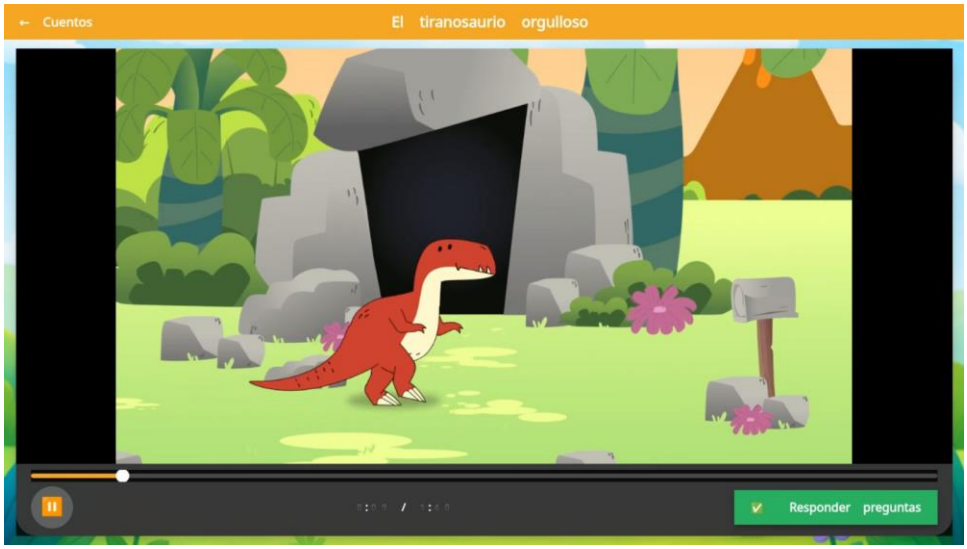


Nota. Elaboración propia.

Al seleccionar cualquiera de los cuentos disponibles, se despliega una ventana dedicada a la reproducción del contenido audiovisual correspondiente, que integra video y audio del cuento elegido, tal como se ilustra en la Figura 32. Una vez finalizada la reproducción del material, el sistema ejecuta de manera automática la transición hacia la ventana de preguntas asociada. Adicionalmente, en la parte inferior de la interfaz se disponen dos botones que permiten al usuario

controlar la navegación: uno de retorno a la pantalla anterior y otro de acceso inmediato a la ventana de preguntas, sin requerir la finalización del video.

Figura 32
Ventana de reproducción de cuentos de la HMI



Nota. Elaboración propia.

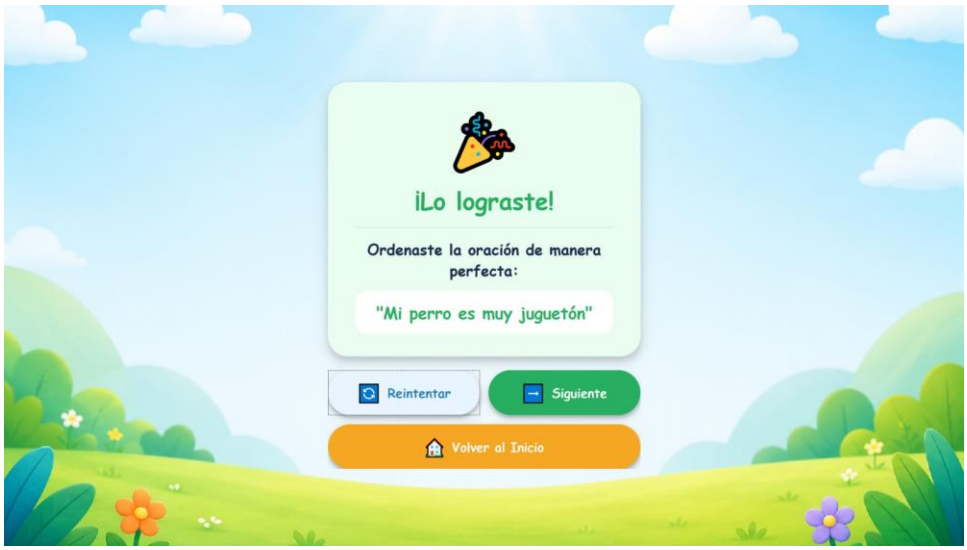
En la Figura 33 se muestra la ventana correspondiente al módulo de preguntas asociado a cada cuento. Esta interfaz integra un conjunto de labels informativos, donde el primero identifica la pantalla como un cuestionario y los subsecuentes presentan las cuatro preguntas que deben ser respondidas. Para cada pregunta se dispone un grupo de tres opciones de respuesta de tipo selección única, restringiendo al usuario a elegir únicamente una alternativa por pregunta. En la parte inferior de la ventana se incorporan dos botones: uno destinado al retorno al menú de cuentos y otro orientado a la validación de las respuestas. Este último ejecuta la revisión del cuestionario y, dependiendo del resultado, despliega una ventana emergente con la retroalimentación correspondiente o habilita el acceso automático a la siguiente ventana del sistema.

Figura 33
Pantalla de preguntas para cuentos

Nota. Elaboración propia.

La pantalla de Felicitaciones, de la Figura 34, aparece al presionar el botón de revisión y tener todas las preguntas correctas, este tendrá un label que felicitará al niño por haber contestado todas las preguntas correctamente y solo poseerá un botón de flecha reversa que nos llevará al menú de selección de juegos

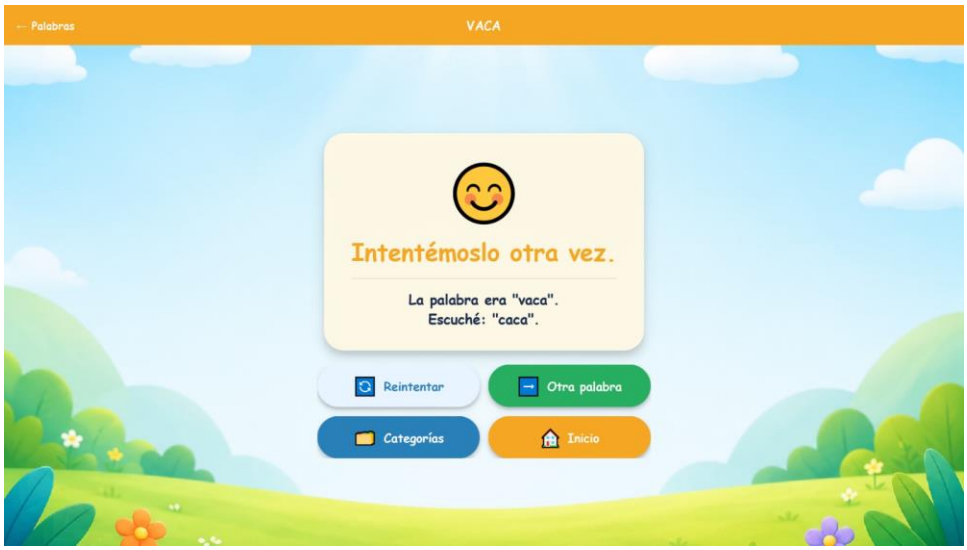
Figura 34
Pantalla de felicitaciones de respuestas correctas.



Nota. Elaboración propia.

La pantalla de Intentémoslo otra vez, de la Figura 35, aparece al presionar el botón de revisión y no tener las respuestas correctas; esta tendrá un mensaje motivador e informativo que le indicará al niño la respuesta esperada junto con la palabra detectada por el sistema, y poseerá botones de acción para reintentar la actividad, cambiar de palabra, regresar al menú de categorías o volver al inicio.

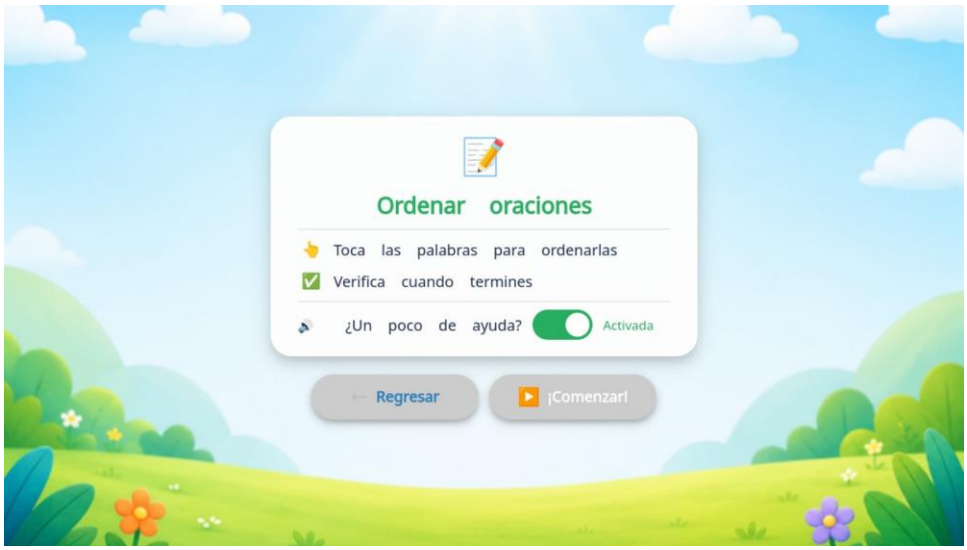
Figura 35
Pantalla de menú de selección de oraciones.



Nota. Elaboración propia.

Al seleccionar el módulo **Ordenar Oraciones** desde el menú de juegos, se despliega la pantalla mostrada en la Figura 36. En esta interfaz se incorpora varios instructivos que indican al usuario la actividad elegida y lo que debe realizar en esta. Debajo de estos elementos se disponen dos botones, uno para regresar a la anterior ventana y otro para empezar con la actividad y acceder a la ventana siguiente del juego.

Figura 36
Pantalla de Ordenar Oraciones.



Nota. Elaboración propia.

La ventana mostrada en la Figura 37 se despliega al seleccionar cualquiera de las oraciones disponibles. En esta interfaz se incluye un label que indica al usuario cómo organizar la oración. Posteriormente, se presenta un área que contiene las palabras desordenadas, así como una zona destinada a su organización mediante un mecanismo de interacción tipo *drag and drop*. En la zona inferior de la pantalla aparecen tres pulsadores que controlan la navegación y la validación del ejercicio: uno para volver a la ventana de selección de oraciones, otro con el que limpiar la zona de organización y uno último para comprobar si efectivamente la oración se ha estructurado correctamente. De acuerdo con el resultado de esta validación el sistema mostrará una ventana emergente con esa retroalimentación o bien hará posible acceder a la pantalla de felicitaciones.

Figura 37
Pantalla de juego de ordenar las oraciones.

Nota. Elaboración propia.

La ventana mostrada en la Figura 38 se despliega al seleccionar el módulo Pronunciación desde el menú principal de juegos. Esta interfaz incorpora un label instructivo que indica al usuario la acción de seleccionar una categoría y una palabra específica de esa categoría.

A continuación, se presenta un conjunto de pushbuttons con representaciones gráficas, los cuales, al ser activados, ejecutan la pronunciación de la palabra correspondiente por parte del sistema y solicitan al usuario su repetición. El sistema evalúa la respuesta fonética emitida y, en función de su nivel de corrección, solicita un nuevo intento o habilita la transición automática hacia la pantalla de felicitaciones. En la parte inferior de la ventana se dispone un pushbutton con iconografía de flecha de retroceso, destinado a retornar al menú de selección de juegos.

Figura 38
Pantalla del juego de Pronunciación.

Nota. Elaboración propia.

El prototipo incorpora, además de los tres juegos anteriores, un cuarto módulo llamado **Platica con el robot**, pensado como un espacio de conversación libre y no como un ejercicio con respuestas correctas o incorrectas. Al ingresar a esta pantalla, el robot saluda al niño con una frase elegida al azar entre varias posibles, mientras un ícono de altavoz con la etiqueta "Escucha" indica que está hablando. Cuando termina el saludo, aparece en la parte inferior un botón con la instrucción "Mantén presionado para hablar": el niño lo mantiene presionado mientras habla y lo suelta cuando termina, de forma parecida a un walkie-talkie.

Figura 39
Pantalla de saludo del módulo Platica con el robot.

Nota. Elaboración propia.

El funcionamiento se basa en el mismo esquema que el juego de pronunciación: durante el tiempo que se mantiene el botón presionado, el sistema registra lo que dice el niño y lo convierte a texto, lo que a su vez se envía al modelo de inteligencia artificial de conversación (la misma tecnología que hay detrás de asistentes como ChatGPT, por ejemplo). El modelo devuelve una respuesta breve (de, como máximo, dos oraciones sencillas) pensada para responder a las explicaciones que le darías a un niño pequeño en español (sin sarcasmos ni contenidos inapropiados), y es el robot el que dicha información por hablar con esta persona antes de volver a escuchar.

Este módulo se conecta a OpenRouter, un servicio que reúne modelos de varios proveedores de inteligencia artificial bajo una sola conexión y que ofrece, entre sus opciones, modelos de uso gratuito. El que usa el prototipo por defecto es Llama 3.3, un modelo de lenguaje grande que la empresa Meta liberó para que cualquiera lo use sin costo. Gracias a eso, la charla con el robot no exige contratar ningún servicio de pago para funcionar, aunque el sistema también

admite conectarse a otros proveedores —como Gemini o Claude— si en el futuro se busca una calidad de respuesta mayor a cambio de una tarifa.

Interacción con los módulos del robot

La GUI funciona enviando señales de estado mediante comunicación serial por cable USB hacia la ESP32, la cual controla:

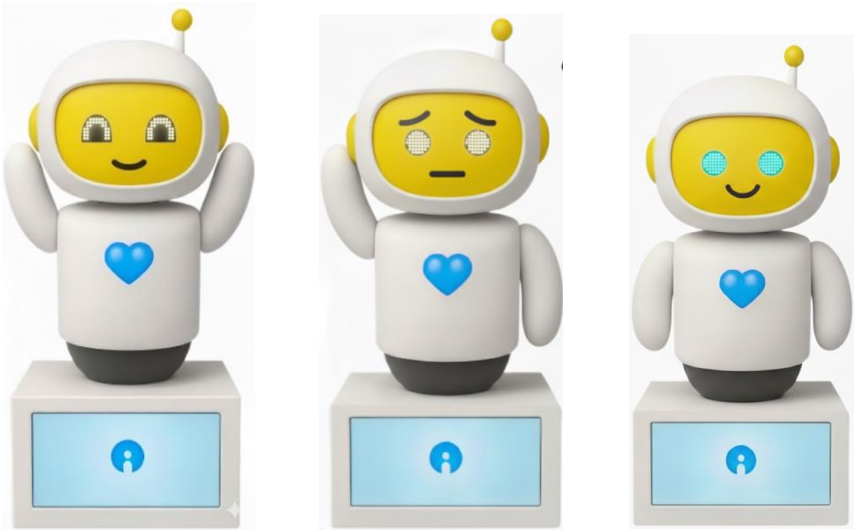
- Servomotores de brazos y cabeza
- Expresiones oculares mediante matrices LED 8×8 con MAX7219

Los tres estados expresivos definidos son:

1. **Estado alegre:** Activado cuando las respuestas son correctas o la pronunciación es adecuada, como se muestra en la Figura 40 (a).
2. **Estado confuso:** Se muestra cuando existen errores en preguntas o pronunciación, observable en la Figura 40 (b).
3. **Estado de espera:** Se activa cuando el robot está escuchando o aguardando la acción del niño, como se visualiza en la Figura 40 (c).

Estos estados acompañan visual y gestualmente la interacción en la interfaz, reforzando la experiencia multisensorial del niño.

Figura 40
Expresiones aproximadas del Robot.



(a) (b) (c)

Nota. Generado por IA.

Arquitectura de comunicación interna

La GUI se integra con los demás módulos mediante una arquitectura distribuida como indica en la Tabla 21.

Tabla 21
Arquitectura de comunicación.

Componente	Protocolo / Método	Función
ASR (reconocimiento de voz)	Selección automática entre Google (en línea) y Vosk (local, sin conexión), según haya Internet disponible	Obtención del texto reconocido y evaluación de la similitud con la palabra esperada
TTS (síntesis de voz)	Motor local conectado a parlantes	Pronunciación de palabras e instrucciones del robot
ESP32 (movimientos y ojos)	Comunicación serial por cable USB (pyserial)	Control de servos y expresiones visuales

Reproductor multimedia	PyQt multimedia / VLC Python	Reproducción de cuentos en video

Nota. Elaboración propia

Esta estructura permite que la interfaz gráfica sea el centro de control pedagógico, mientras que la ESP32 es el centro de control motriz del robot.

Fundamentación del diseño

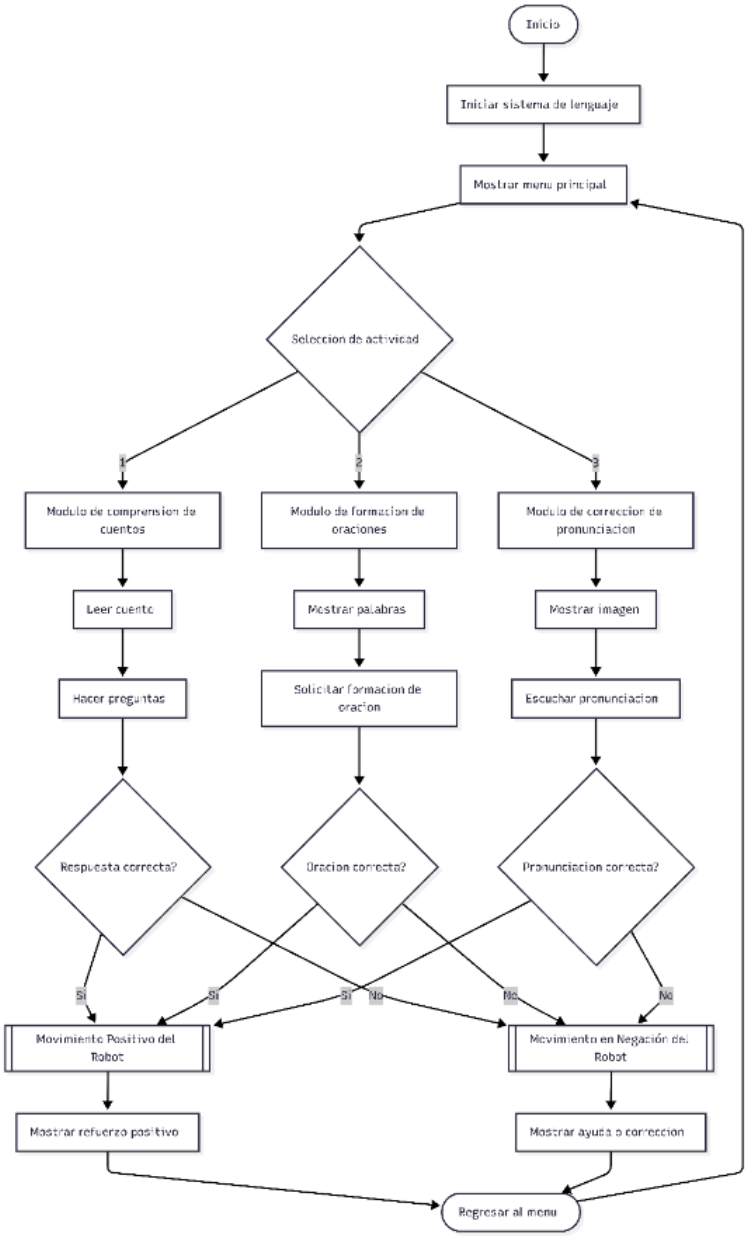
La arquitectura se diseñó considerando:

- La necesidad de **interacción clara y directa** para niños de 4 a 5 años.
- La separación estricta entre:
    - Presentación (ventanas Qt),
    - Procesamiento de voz.
    - Controladores
    - Hardware y módulos
- Optimización para el rendimiento de la Raspberry Pi 5.
- Reducción del tiempo de carga y fluidez en animaciones y transiciones.
- Minimización de errores y robustez ante entradas inesperadas del usuario infantil.

Desarrollo de Software

Se procedió con la implementación de la programación siguiendo el esquema mostrado a continuación:

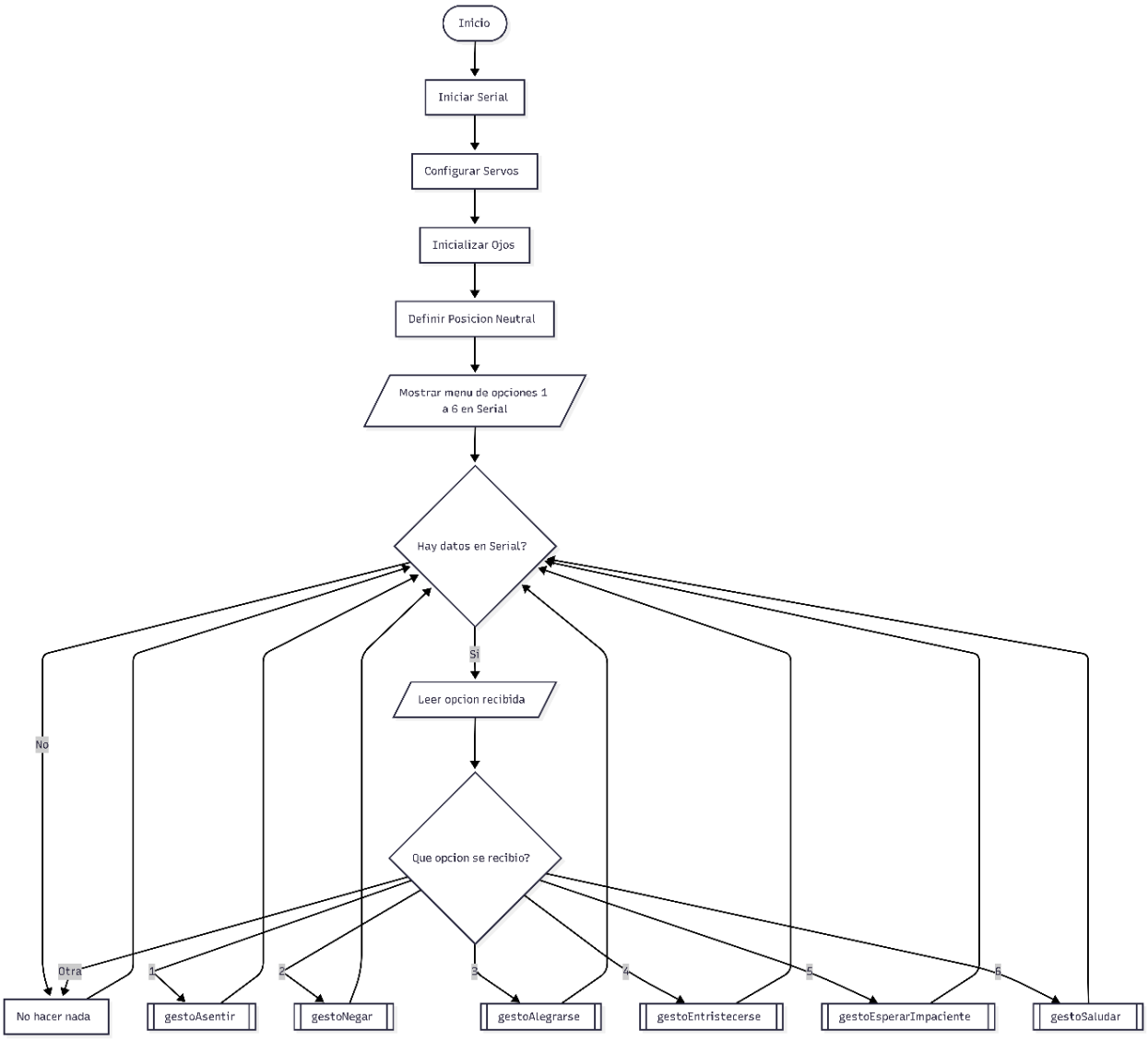
Figura 41
Prueba de funcionamiento mecánico preliminar.



Nota: Elaboración propia.

Esta programación representa el cerebro principal del sistema controlando todo el proceso a través de la raspberry pi, además se comunica con arduino para orquestar los movimientos del robot según la solicitud principal de la siguiente forma:

Figura 42
Prueba de funcionamiento mecánico preliminar.



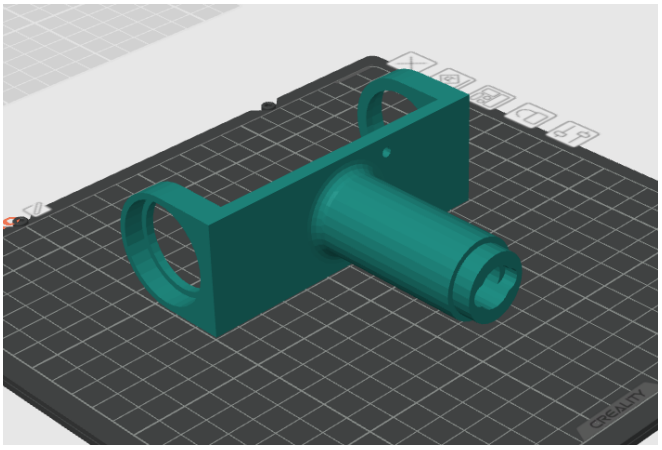
Nota: Elaboración propia.

Fabricación y ensamblaje del prototipo

Tomando en cuenta el método de fabricación seleccionado, correspondiente a la tecnología de manufactura aditiva mediante impresión 3D por deposición de material fundido (FDM), los diseños de las piezas fueron desarrollados con el objetivo de minimizar la necesidad de generación de soportes durante la impresión, con el fin de disminuir el consumo de material destinado a elementos que no forman parte de la pieza final y que son considerados desecho no aprovechable. Posteriormente, los modelos fueron exportados desde el entorno CAD del software Inventor al formato STL, el cual es requerido por las herramientas de laminado para la impresión 3D. Una vez generados los archivos, estos fueron importados al software Orca Slicer, donde se realizó la configuración de los parámetros necesarios para el proceso de impresión.

Para la configuración principal se tomaron en cuenta 2 tipos de piezas, las piezas internas que soportan los esfuerzos principales y las piezas externas que marcan la forman o estética del robot, empezando por las piezas internas, aquella que requiere más atención es el eje principal del cuello, que soporta todo el sistema de movimiento de la cabeza.

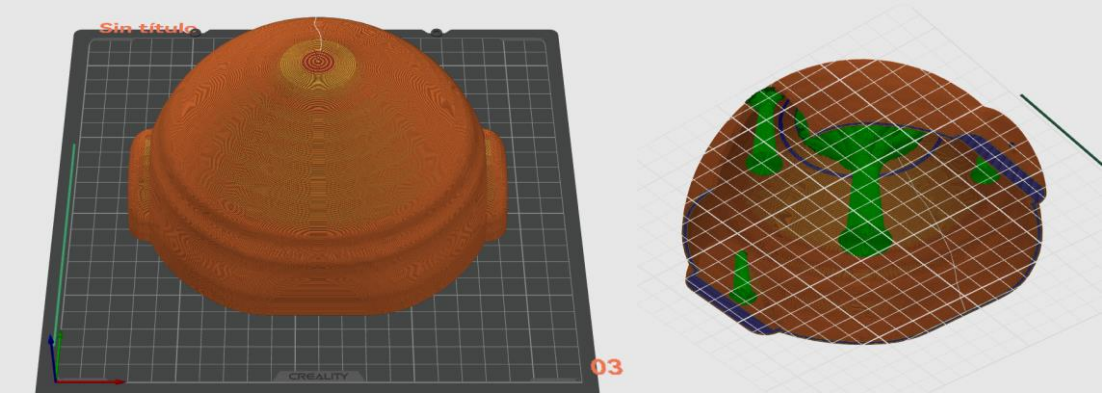
Figura 43
Posición de impresión del eje principal del cuello.



Nota: Elaboración propia.

La importancia de la configuración de este modelo mostrado en la Figura 43 radica en que la impresión 3D no genera un modelo isotrópico, pues la impresora deposita el material en la base y sube por capas, de manera que las fibras se generan de manera horizontal a la base, siendo este plano el más resistente, por lo cual se asentó el modelo en la cama de impresiónpara que este pueda soportar mejor los esfuerzos de tensión y torsión generados por el movimiento y peso de la cabeza, se mantuvo una velocidad nominal de 90 mm/s correspondiente a lo recomendado por el fabricante de la impresora, dos perímetros sólidos y un relleno del 10% manteniendo suficiente firmeza sin ganar demasiado peso, en cuanto a la altura de capa se mantuvo el valor estándar de 0.2mm.

Figura 44
Configuración del STL de la cabeza.



Nota: Elaboración propia.

Para el modelo mostrado en la Figura 44, al corresponder a una parte externa que va a ser visible, se optó por reducir la altura de capa a 0.12mm para aumentar la definición de las curvas, reducir la velocidad a un 60% de la nominal alrededor de 60mm/s con el fin de mejorar el acabado de las superficies, y la posición en la placa base está dada para mantener los soportes en la parte interna, evitando desperfectos en la parte externa del mismo, además se aumentó el

número de perímetros sólidos a 3, con el objetivo de tener más rango de trabajo para el procesamiento posterior.

Con toda la configuración necesaria se procedió a realizar la impresión de manera normal, exportando el archivo con el código G a la memoria de la impresora 3d, una vez impresos todos los elementos ser retiró los soportes de aquellos que lo tenía, y se procedió a la prueba inicial, mediante una prueba real de funcionamiento mecánico, se juntó las piezas con los componentes y se comprueba que todo encaje como fue diseñado, obteniendo el primer vistazo del robot como se visualiza en la Figura 45.

Figura 45
Prueba de funcionamiento mecánico preliminar.

Nota: Elaboración propia.

Con todos estos elementos juntos se realizó la primera prueba mecánica corroborando que todos elementos están correctamente impresos para su ensamblaje y el movimiento mecánico se realiza de acuerdo a los parámetros esperados, con esto en mente se procede a

realizar el procesamiento del acabado de las piezas, por el proceso de lijado se retiraron las posibles imperfecciones existentes en las partes externas de la misma, dejando la superficie lo más uniforme posible, para posteriormente agregar una capa de fondo gris o primer, que cubre toda la pieza dándole un acabado más liso como se muestra en Figura 46

Figura 46
Post-procesado del robot.

Nota: Elaboración propia.

Posteriormente pasar al proceso de pintura, donde se adhiere el color metalizado que le da el brillo necesario para cumplir con el estándar de color solicitado, tal y como se muestra en la Figura 46.

Figura 47
Diseño final del robot.

Nota: Elaboración propia.

Reconocimiento de voz

El reconocimiento automático de voz del prototipo funciona mediante un sistema híbrido que combina dos motores de reconocimiento distintos: uno en línea, que utiliza el servicio de Google a través de Internet, y uno local, basado en el motor de código abierto Vosk, que trabaja completamente sin conexión. Antes de cada intento de reconocimiento, el sistema comprueba de forma automática si existe conexión a Internet disponible en ese momento. Si la hay, se utiliza el motor en línea de Google, que ofrece un reconocimiento más preciso; si no la hay, el sistema recurre de manera transparente al motor local Vosk, que aunque es algo menos

exacto, permite que el robot siga funcionando incluso en lugares sin acceso a Internet, como suele ocurrir en muchas instituciones educativas. No se optó por usar un único motor porque cada uno tiene una ventaja distinta que conviene aprovechar según el contexto: el motor en línea da mejores resultados cuando hay conectividad, mientras que el motor local garantiza que el robot nunca deje de funcionar por falta de Internet. Combinar ambos permite obtener la mejor precisión posible sin sacrificar la disponibilidad del sistema, y el cambio entre uno y otro ocurre de forma automática, sin que el usuario tenga que intervenir ni darse cuenta de cuál de los dos se está usando en cada momento. Estas características hacen que la solución sea especialmente adecuada para la aplicación educativa, pensada para funcionar de forma confiable tanto en instituciones con buena conectividad como en aquellas que no la tienen.

El reconocimiento automático de voz busca convertir una señal de audio producida por un usuario en una representación textual que pueda ser procesada por el software que controla el robot. Este proceso implementa diversas etapas de procesamiento digital y análisis estadístico mediante las cuales el sistema es capaz de identificar los sonidos que ha pronunciado y de determinar la secuencia de palabras que tiene una alta probabilidad de corresponder al habla del usuario.

En primer lugar, la voz del niño es grabada por medio de un micrófono enlazado a la Raspberry Pi. La señal acústica que el niño produce al pronunciar las palabras es transformada por el micrófono en una señal eléctrica que, posteriormente, es digitalizada mediante un conversor analógico-digital, convirtiéndose en una secuencia de muestras digitales que describen la forma de onda del audio y que serán la entrada del sistema de reconocimiento de voz. A continuación, la señal de audio entra a una fase de preprocesado donde el flujo de muestras es dividido en pequeños segmentos temporales denominados frames; sobre cada uno de estos frames se aplican técnicas de procesamiento digital de señal que permiten

resaltar las características relevantes del habla y reducir el efecto del ruido ambiental, o de otras perturbaciones que puedan existir en el momento de la grabación del audio. Por último, el sistema realiza la extracción de una representación de la señal de voz a partir de un conjunto de características acústicas, que permite identificar las propiedades fonéticas de cada uno de los sonidos pronunciados, reduciendo la cantidad de información que es necesaria para el reconocimiento sin perder las características que permiten diferenciar unas palabras de otras.

Dentro de ese preprocesado, el prototipo incorpora un mecanismo propio para decidir en qué momento debe detenerse la grabación, que combina dos criterios a la vez. Por un lado, aplica un filtro que elimina las frecuencias muy bajas correspondientes al zumbido eléctrico típico de la red doméstica, entre 50 y 60 Hz, que en pruebas con el micrófono real del robot llegó a confundirse con la voz del niño e impedía detectar correctamente los silencios. Por otra parte, el sistema determina de manera continúo un nivel de referencia del fondo de ruido del entorno a la vez que graba, en lugar de usar un valor fijo predefinido, de modo que se adapta automáticamente tanto en salones silenciosos como en salas con más ruido del entorno. La grabación parará cuando el nivel de nivel de audio caiga lo suficiente con respecto al nivel más alto detectado de la grabación en esa sesión, o cuando se acerque lo suficiente al nivel de fondo de ruido estimado en ese momento, lo primero que suceda. Este doble criterio permite que la detección de silencio funcione de forma confiable en distintos entornos, sin necesidad de recalibrar manualmente el sistema cada vez que cambian las condiciones del lugar donde se usa el robot.

Una vez que se han obtenido las características acústicas, en el caso de estar usando el motor local Vosk, éste utiliza modelos acústicos que han sido entrenados con técnicas de aprendizaje automático para así estimar la secuencia de fonemas generados en la señal de entrada. Después, un modelo lingüístico se encarga de ponderar entre las diferentes

combinaciones de palabras y de forma explícita tiene en cuenta la gramática del idioma, dando como resultado la hipótesis de reconocimiento con la mayor probabilidad. Al final de este proceso, el sistema genera el texto que representa la pronunciación hecha por el usuario. Cuando, en cambio, se está usando el motor en línea de Google, este mismo tipo de procesamiento ocurre en los servidores de Google en lugar de en la Raspberry Pi, y el resultado, el texto reconocido, es lo único que se recibe de vuelta.

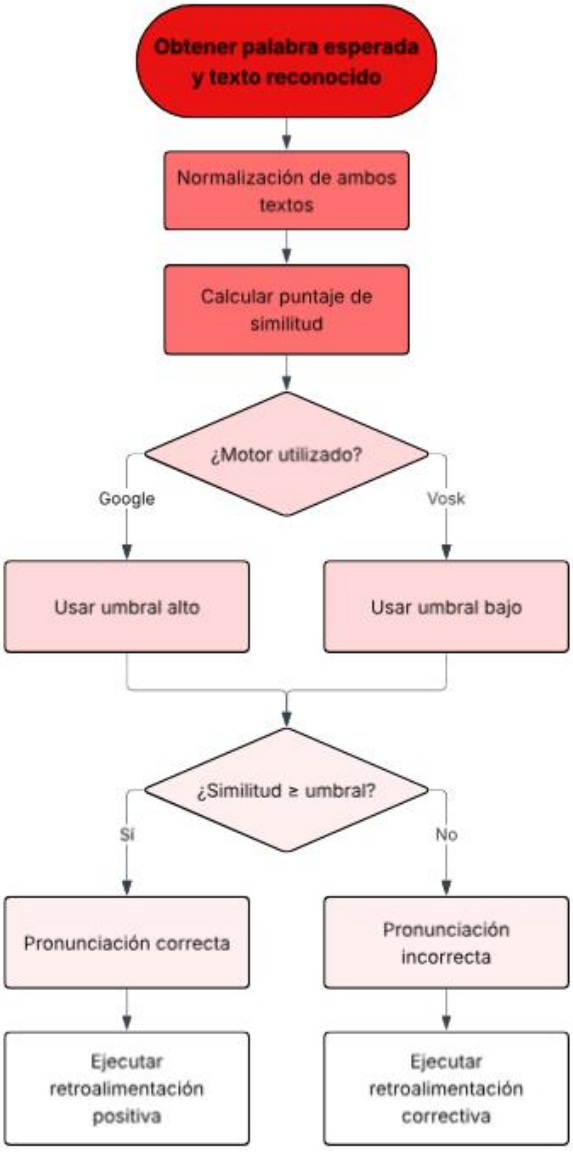
El flujo general seguido por el sistema de reconocimiento de voz durante este proceso está en la Figura 48, donde aparecen las distintas fases que van desde la captura de audio hasta la obtención del texto del motor de reconocimiento de voz.

Figura 48
Resumen del proceso de reconocimiento de voz.

Nota: Elaboración propia.

Para determinar si la pronunciación del niño fue correcta, el sistema no exige una coincidencia exacta entre el texto reconocido y la palabra esperada, ya que esto sería demasiado estricto y podría penalizar diferencias mínimas de transcripción que no reflejan necesariamente un error real de pronunciación. En su lugar, primero se normaliza la palabra esperada y el texto reconocido, convirtiéndolos a minúsculas, eliminando tildes y símbolos solamente se dejan letras y números; y se calcula un puntaje de similitud, de acuerdo a la cantidad de cambios que se le tienen que hacer a una palabra para convertirla en la otra palabra. Cuanto menor es la cantidad de cambios respecto a la longitud de las palabras que se comparan, mayor es el puntaje de similitud, el cual puede ir de 0 (sin ningún tipo de parecido) a 1 (son exactamente iguales). La pronunciación se considera correcta cuando ese puntaje supera un umbral mínimo, el cual varía según el motor de reconocimiento que se haya usado en ese intento: al ser el motor en línea de Google más preciso, se exige un puntaje más alto para aceptar la respuesta; mientras que, al ser el motor local Vosk menos preciso —y tender a confundir ciertos sonidos parecidos entre sí, como ocurre por ejemplo entre "f" y "p", se exige un puntaje más bajo, para no penalizar al niño por errores que en realidad son del motor de reconocimiento y no de su pronunciación. Por ejemplo, si el niño pronuncia la palabra "manzana" y el motor de reconocimiento genera un texto muy parecido, aunque no sea idéntico letra por letra, el sistema puede considerar igualmente la respuesta correcta cuando el puntaje de similitud supera el umbral correspondiente, y ejecuta las acciones de retroalimentación definidas. Este proceso se representa en la Figura 49.

Figura 49
Proceso de pronunciación.



Nota: Elaboración propia.

Este mecanismo permite implementar ejercicios interactivos de pronunciación y reconocimiento verbal, favoreciendo el desarrollo de habilidades lingüísticas y la retroalimentación inmediata durante el proceso de aprendizaje infantil.

Captura de audio

La captura de audio del prototipo se lleva a cabo a partir de la librería sounddevice, una interfaz que se ha desarrollado para Python y que permite acceder a los dispositivos de entrada y salida de audio del sistema operativo, utilizando para ello la biblioteca PortAudio. Su función dentro del prototipo consiste en contrarrestar un canal de comunicación entre el micrófono USB y la aplicación principal ejecutada en la Raspberry Pi, permitiendo realizar una adquisición ininterrumpida de la señal de voz emitida por el usuario, dada las actividades educativas.

El funcionamiento de sounddevice se basa en la creación de un flujo audio (audio stream), a partir del cual las muestras capturadas por el micrófono son recibidas continuamente, conforme son adquiridas por el dispositivo. En lugar de esperar a que se termine una grabación completa, la librería divide el flujo de audio en pequeños bloques de datos que son entregados de forma inmediata a la aplicación. Este mecanismo de adquisición ininterrumpida reduce la latencia del sistema y permite que el procesamiento del audio se realice prácticamente en tiempo real.

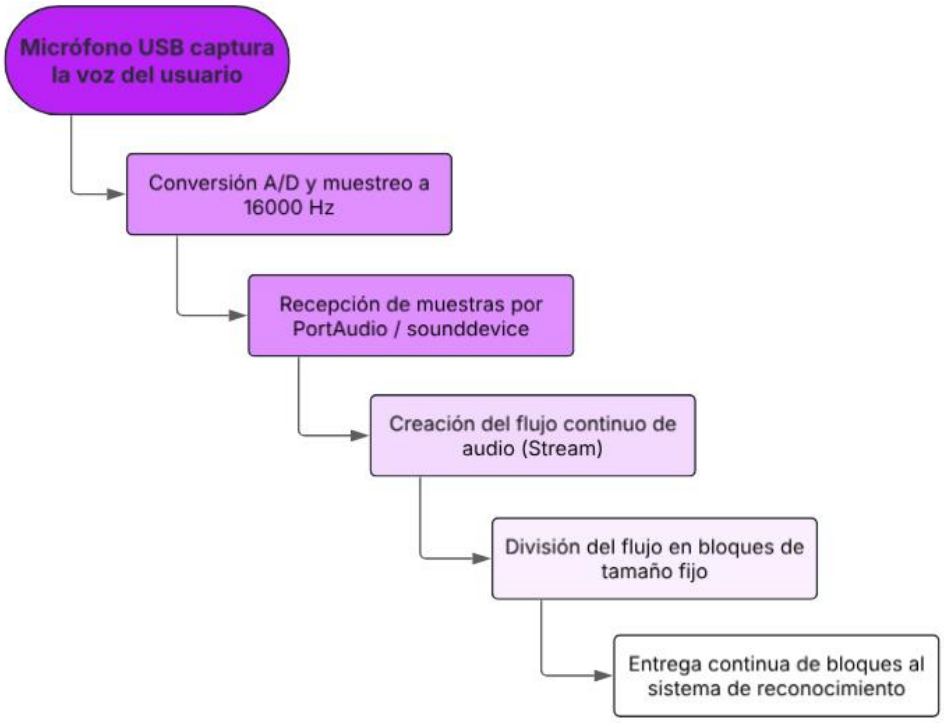
El dispositivo de adquisición convierte, en el momento de la captura, la señal analógica del micrófono en una secuencia de muestras digitales que se transmiten a la Raspberry Pi. En esa computadora, la biblioteca sounddevice recoge las muestras y las coloca en bloques de tamaño fijo que van directamente al motor de reconocimiento de voz. Este trabajo por bloques genera un flujo continuo de información entre el dispositivo de captura y el motor de reconocimiento, de modo que no hay interrupciones durante la interacción con el usuario.

El sistema utiliza para su funcionamiento una frecuencia de muestreo de 16000 Hz, lo que es muy conveniente porque 16 kHz es un valor utilizado en automático debido a que ofrece un compromiso entre la calidad de la señal y su reconocimiento sin un consumo excesivo de recursos. Como la mayor parte de la información pertinente se encuentra en frecuencias por

debajo de los 8.000 Hz, el muestreo a 16 kHz se ajusta al requerimiento del teorema de Nyquist, que establece que una frecuencia de muestreo superior a la de 2 veces la de la frecuencia del contenido permite una representación correcta de las componentes espectrales de la voz y que, a la vez, no genera un exceso de datos en el sistema.

El esquema general seguido durante la captura y el envío del audio hacia el motor de reconocimiento de voz se encuentra ilustrado en la Figura 50.

Figura 50
Proceso de pronunciación incorrecta.



Nota: Elaboración propia.

La integración entre sounddevice y los motores de reconocimiento de voz permite capturar el audio de manera continua y con baja latencia, sin importar cuál de los dos motores esté activo en cada momento. Cuando el sistema opera con el motor local Vosk, todo el proceso

de captura, transmisión y procesamiento del audio se realiza dentro de la propia Raspberry Pi, sin depender de servicios externos, lo cual es especialmente valioso en entornos educativos sin conexión a Internet. Cuando, en cambio, hay conexión disponible y se utiliza el motor en línea de Google, sounddevice cumple la misma función de captura continua, y únicamente el audio ya grabado se envía a través de Internet para su reconocimiento, de modo que la experiencia de uso resulta similar para el niño sin importar cuál motor esté funcionando en ese momento.

Generación de voz

El prototipo utiliza Piper, un motor neuronal de código abierto diseñado para una eficiencia óptima en dispositivos con recursos restringidos, para la síntesis de voz, como la Raspberry Pi. Los sistemas antiguos generaban la voz concatenando fragmentos de audio previamente grabados; Piper, en cambio, usa un modelo neuronal entrenado de antemano que aprende las características acústicas y lingüísticas del habla a partir de grandes conjuntos de datos, y logra así una pronunciación fluida y una entonación cercana a la de una voz humana.

El modelo de voz se carga una sola vez al iniciar el sistema y se mantiene disponible durante toda la ejecución del programa, de modo que no hace falta volver a cargarlo cada vez que el robot necesita hablar. Para el español se usa un modelo de voz femenina entrenado específicamente para ese idioma, aunque el sistema permite intercambiar este modelo por otro disponible sin reiniciar la aplicación, lo que facilita adaptar la voz del robot según las preferencias de quien lo configure.

El funcionamiento general del modelo empieza por el procesamiento del texto de entrada. En esta etapa se interpreta la secuencia de caracteres y la estructura del lenguaje, de manera que el sistema pueda determinar la pronunciación adecuada de cada palabra y una entonación coherente para toda la frase. A partir de esa información, el modelo neuronal genera directamente la señal de audio correspondiente.

Otros enfoques requieren generar y almacenar con anticipación todos los audios que el robot podría necesitar; en este prototipo, en cambio, la síntesis de voz ocurre en tiempo real. Cada vez que el robot necesita decir una frase, el modelo genera el audio en pequeños fragmentos que se reproducen de manera prácticamente inmediata por el parlante, sin esperar a que se complete toda la síntesis antes de empezar a escucharse. Esto es posible porque Piper es un modelo bastante más liviano que otros sistemas de síntesis neuronal de mayor tamaño, lo que le permite ejecutarse directamente en la Raspberry Pi sin requerir una unidad de procesamiento gráfico (GPU) ni generar tiempos de espera perceptibles para quien escucha, a pesar de los recursos limitados del hardware embebido del prototipo.

El sistema incluye además algunas funciones adicionales pensadas para el uso educativo del robot. La ampliación de la voz puede ser controlada de forma independiente al volumen general del sistema, que hace que el robot sea controlado con un volumen suficiente y no solo en función de la configuración de audio del dispositivo. También se puede ajustar la velocidad a la que se pronuncia una frase. Esta funcionalidad, por ejemplo, se utiliza en el momento en que el robot le pide al niño que repita una palabra: primero se dice la instrucción de forma normal y, tras una breve pausa, se dice la palabra objetivo de forma más lenta, de manera que el niño pueda escuchar la palabra objetivo con suficiente claridad antes de intentar repetirla.

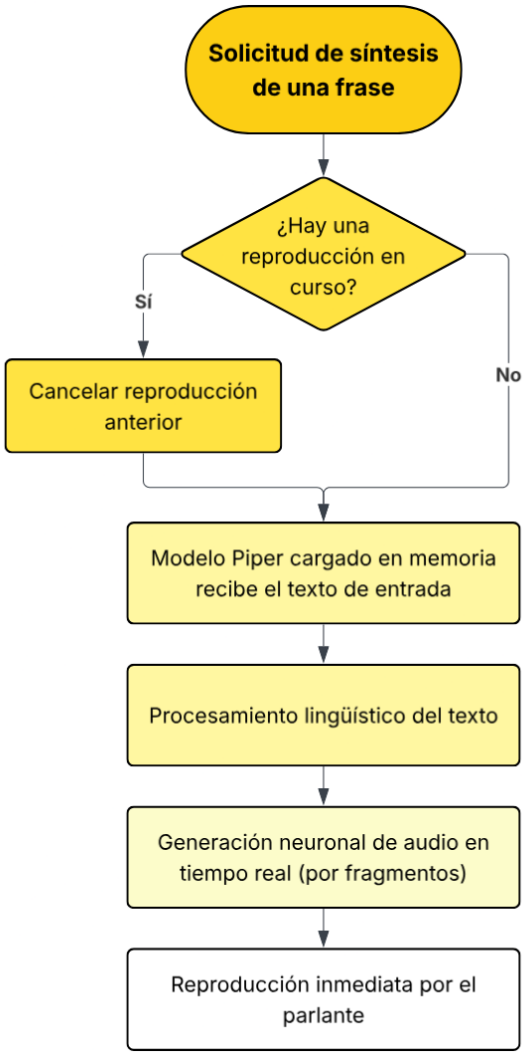
Por último, en cualquier momento en que se solicite una nueva frase, la reproducción de la frase anterior se cancela de inmediato, de manera que el robot siempre esté diciendo únicamente lo último que se le pidió, sin acumular ni superponer distintas reproducciones. Esto hace que la interacción con el niño se sienta más natural y con una respuesta más inmediata.

En conjunto, esta estrategia de síntesis de voz en tiempo real permite que el robot cuente con una voz natural para cualquier frase que necesite decir, sin las limitaciones de un conjunto

fijo de audios pregrabados, y sin exigir al hardware embebido una capacidad de cómputo mayor a la que puede ofrecer razonablemente una Raspberry Pi.

Figura 51

Resumen del Reconocimiento de Voz



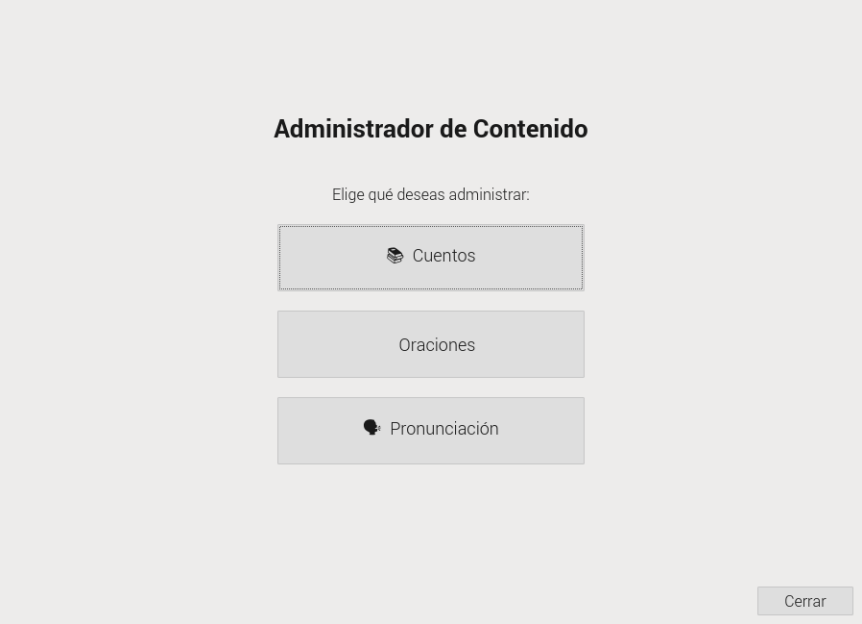
Nota: Elaboración propia.

Administrador de Contenido

Más allá de las tres actividades que ve el niño, el proyecto incluye una herramienta aparte pensada para la profesora: el Administrador de Contenido. Es un programa independiente que

se abre en la computadora, no en el robot, y permite agregar, editar o eliminar cuentos, oraciones y palabras de pronunciación sin necesidad de abrir ningún archivo ni de saber programar. Su razón de ser es sencilla: el contenido de un robot educativo se queda corto muy rápido si siempre son los mismos diez cuentos y las mismas cien palabras, y no tendría sentido depender de un programador cada vez que la institución quiera renovar el material o adaptarlo a un grupo distinto de niños.

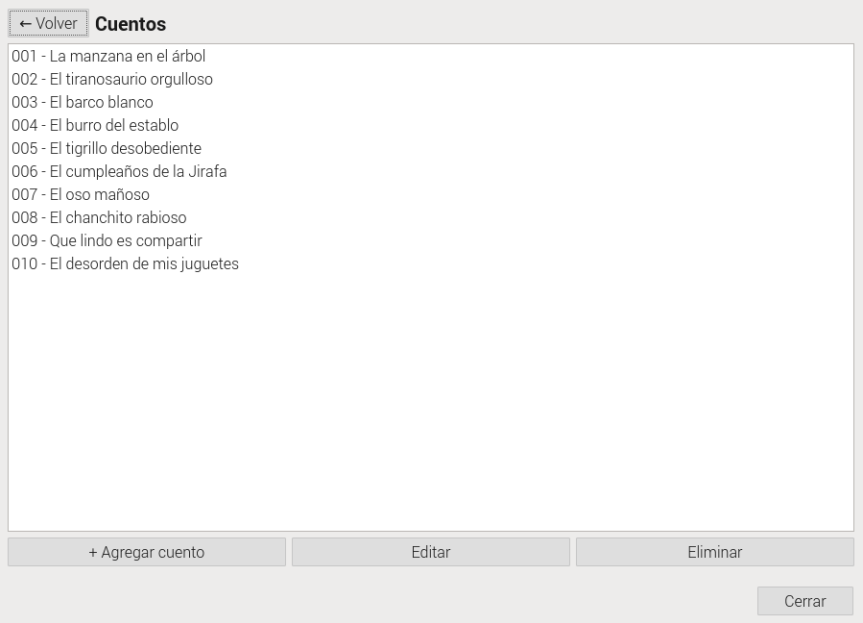
Figura 52
Pantalla de inicio del Administrador de Contenido.



Nota. Elaboración propia.

Desde la pantalla de inicio, la profesora elige qué quiere administrar. En Cuentos aparece la lista completa de historias cargadas, cada una con su número y su título, y tres botones simples para agregar uno nuevo, editar el seleccionado o eliminarlo.

Figura 53
Pantalla de administración de Cuentos.



Nota. Elaboración propia.

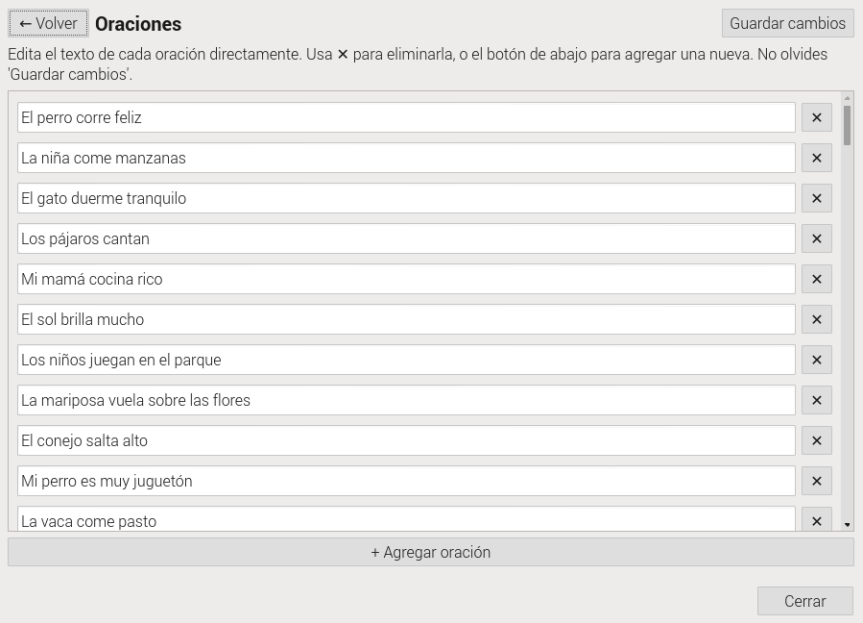
Al editar un cuento se abre un formulario donde se puede cambiar el título, elegir una nueva imagen de portada o un nuevo video con el explorador de archivos del propio computador, y administrar las preguntas de comprensión asociadas a esa historia —agregarlas, editarlas o quitarlas una por una—. El programa recorta y ajusta automáticamente cualquier imagen que se seleccione, así que la profesora no tiene que preocuparse por el tamaño o el formato correcto: solo elige la foto que quiere usar.

Figura 54
Formulario de edición de un cuento.

Nota. Elaboración propia.

La pantalla de Oraciones sigue la misma filosofía, pero es todavía más simple: cada oración es una casilla de texto que se edita directamente, con una ✕ al lado para borrarla y un botón para agregar una nueva al final. El programa separa la oración en palabras automáticamente, sin que la profesora tenga que preocuparse por ese detalle —solo escribe la oración completa como la diría un niño, y el sistema arma con ella el ejercicio de ordenar palabras—. Si se sale de la pantalla con cambios sin guardar, el programa avisa antes de perderlos.

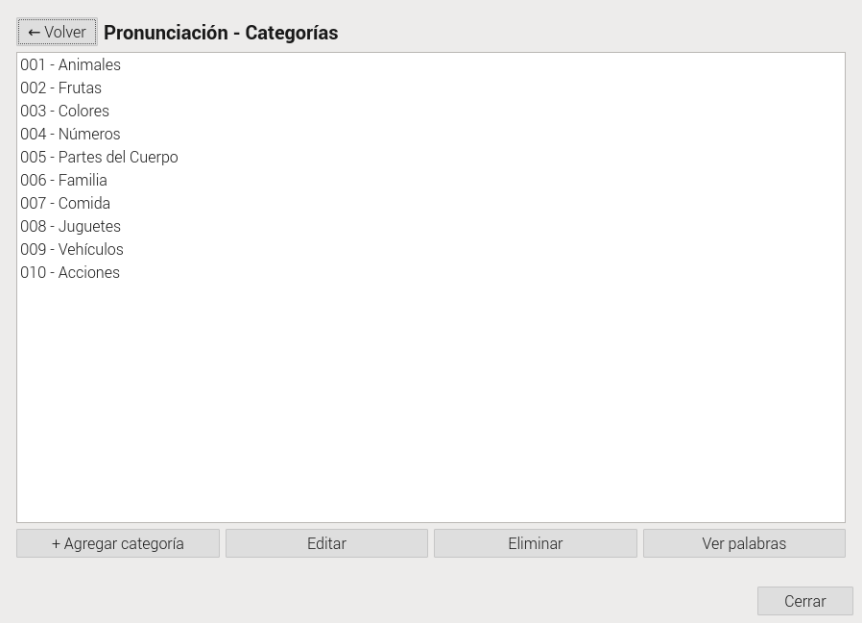
Figura 55
Pantalla de administración de Oraciones.



Nota. Elaboración propia.

En Pronunciación, la lógica es parecida a la de Cuentos: se administran categorías completas —Animales, Frutas, Colores, y así sucesivamente—, y dentro de cada una, las palabras y sus imágenes asociadas.

Figura 56
Pantalla de administración de categorías de Pronunciación.



Nota. Elaboración propia.

En conjunto, el Administrador de Contenido resuelve un problema práctico de cualquier robot educativo: mantener el material actualizado y relevante sin depender de conocimientos técnicos. Una profesora puede cambiar los cuentos según la temática del mes, agregar palabras de pronunciación relacionadas con lo que se está enseñando en clase, o adaptar las oraciones al nivel del grupo, todo desde una interfaz sencilla con botones claros y mensajes de confirmación antes de guardar o borrar algo. Esta posibilidad de personalización fue, de hecho, uno de los aspectos mejor calificados en las encuestas de validación descritas en el Capítulo IV, donde tanto los docentes de Educación Inicial como los de Ingeniería puntuaron el módulo de gestión de contenido entre los más altos de sus respectivas encuestas.

Capítulo IV

Pruebas y resultados

Este capítulo muestra las pruebas de funcionamiento que se han llevado a cabo en el robot educativo diseñado para validar el correcto funcionamiento de los diferentes subsistemas que articulan el prototipo. Para ello se analizan, en las pruebas, el funcionamiento mecánico de las juntas, la realización de las expresiones emocionales, la interacción entre la parte electrónica y el software, así como el funcionamiento de las funciones implementadas para el reconocimiento de la voz, para el reconocimiento de las imágenes y las actividades educativas implementadas. En cada prueba se han diseñado procedimientos particulares y los criterios de evaluación que permitieron analizar la precisión, estabilidad y fiabilidad del sistema verificando que el robot cumplía con los requerimientos que se establecían durante el desarrollo del proyecto, demostrando así su viabilidad como herramienta para el aprendizaje del lenguaje en niños de 4 a 5 años.

Pruebas del Sistema Mecánico

Esta sección presenta las pruebas realizadas al sistema mecánico del robot, orientadas a verificar la precisión angular de los movimientos de brazos y cabeza al ejecutar dos expresiones gestuales: felicidad y tristeza. Para cada expresión se compararon los ángulos programados con los ángulos realmente alcanzados por los servomotores, con el fin de determinar el error angular y la repetibilidad del movimiento.

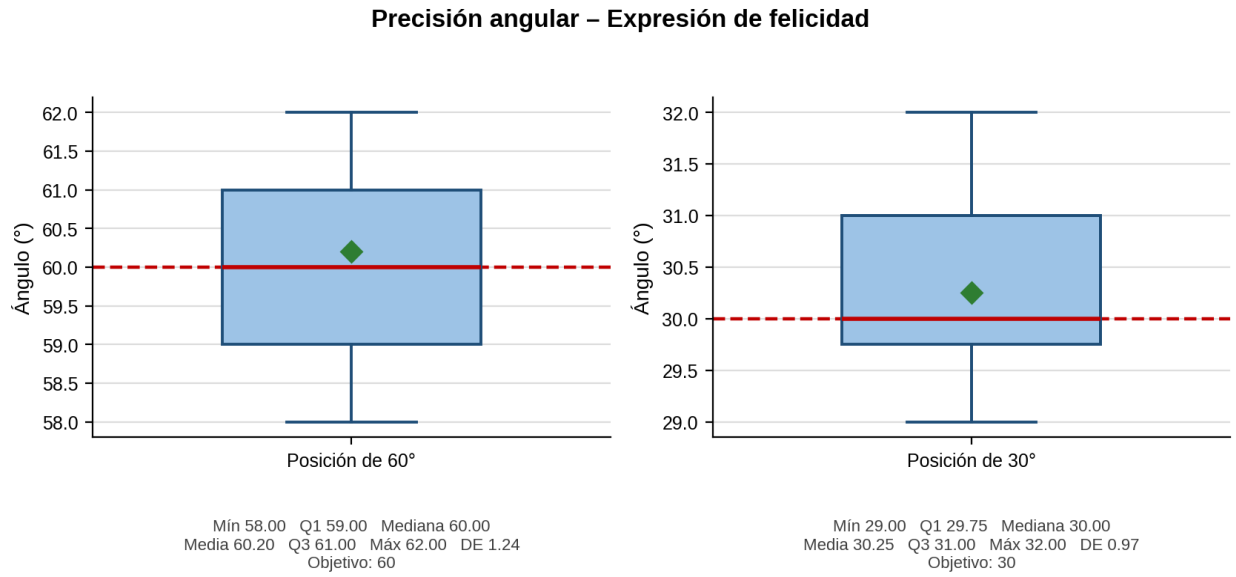
Pruebas de precisión angular en: Expresión de felicidad

Movimientos evaluados:

- Posición inicial: 0° (brazos abajo)
- Movimiento 1: 60°
- Movimiento 2: 30°

- (Repitiendo cuatro veces el movimiento 1 y 2)
- Posición final: 0°
- La cabeza permanece inmóvil

Figura 57
Pruebas en la expresión de felicidad



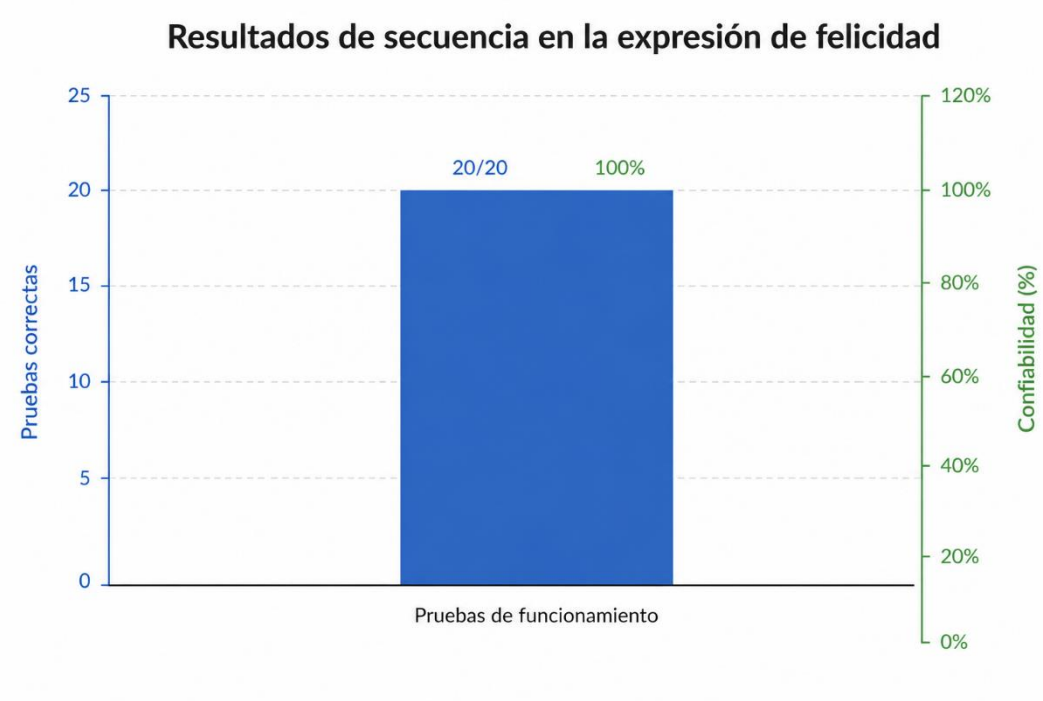
Nota. Elaboración propia.

Tabla 22
Resumen de resultados de movimiento en la expresión de felicidad.

Parámetro	Número de pruebas	Promedio	Error absoluto promedio	Error porcentual
Posición de 60°	20	60.2°	1°	0.75°
Posición de 30°	20	30.25°	1.67%	2.5%

Nota. Elaboración propia

Figura 58
Resumen de resultados en la secuencia de expresión de felicidad



Nota. Elaboración propia.

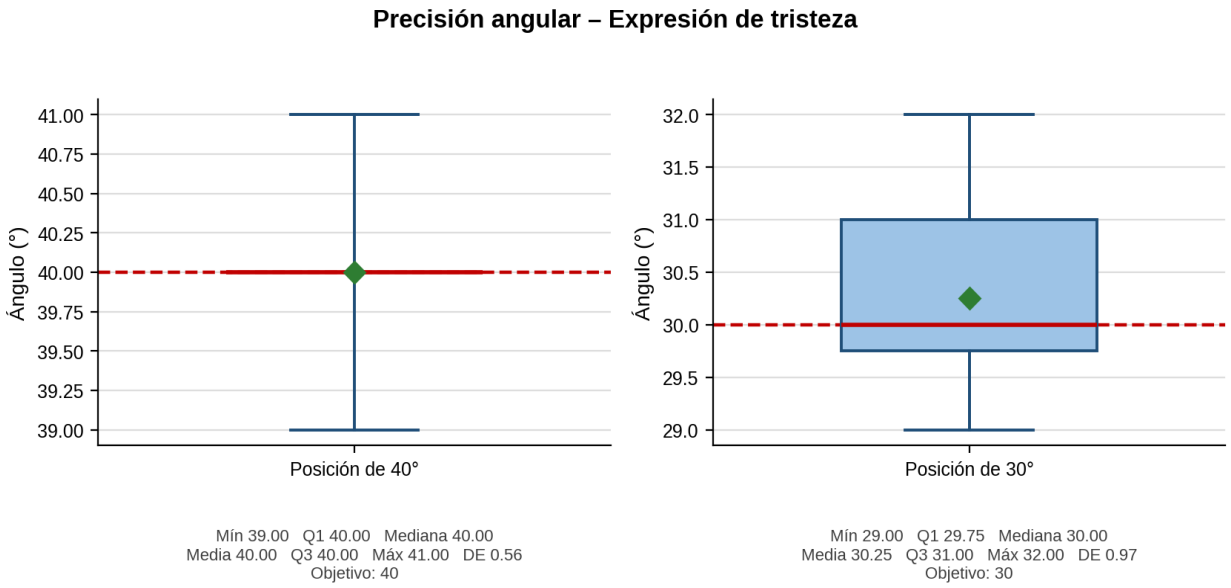
Pruebas de precisión angular en: Expresión de tristeza

Movimientos evaluados:

- Brazos: 0° → 40°
- Cabeza: 0° → 30°

Figura 59

Pruebas en la expresión de tristeza



Nota. Elaboración propia.

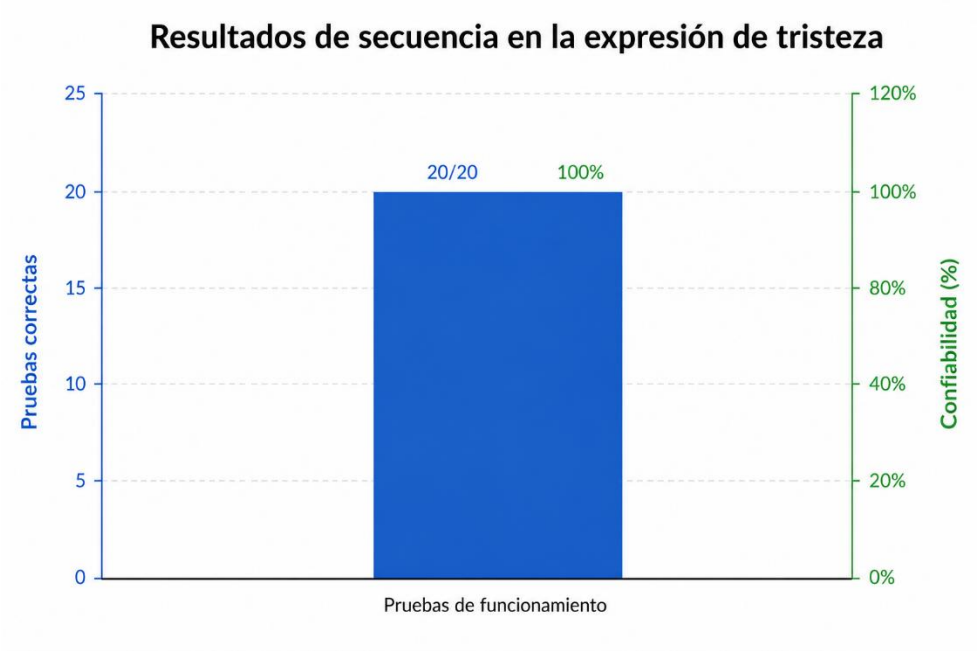
Tabla 23

Resumen de resultados de movimiento en la expresión de tristeza.

Parámetro	Número de pruebas	Promedio	Error absoluto promedio	Error porcentual
Posición de 40°	20	40°	0.3°	0.75%
Posición de 30°	20	30.25°	0.75°	2.5%

Nota. Elaboración propia

Figura 60
Resumen de resultados en la secuencia de expresión de tristeza.

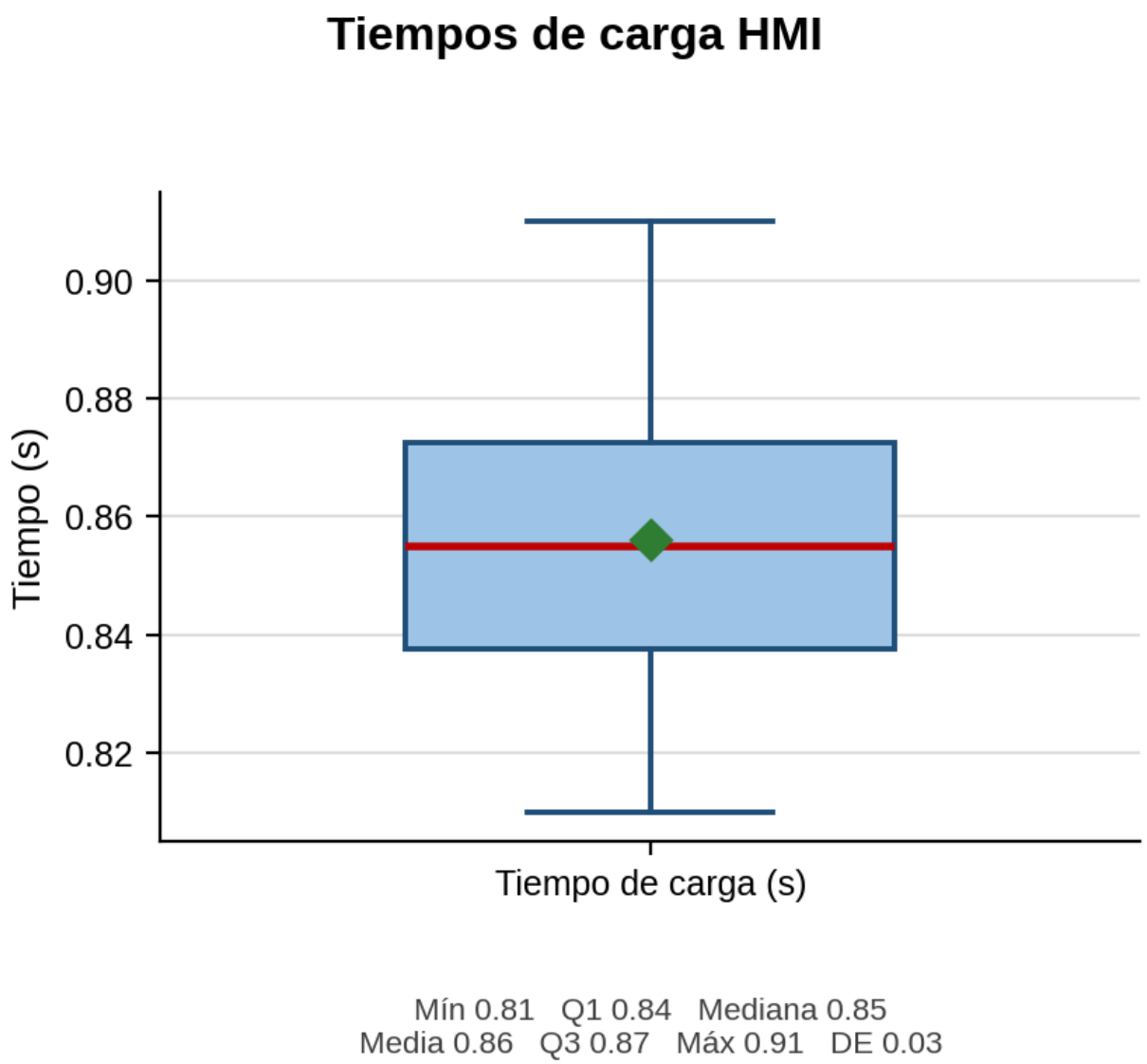


Nota. Elaboración propia.

Pruebas de la interfaz gráfica

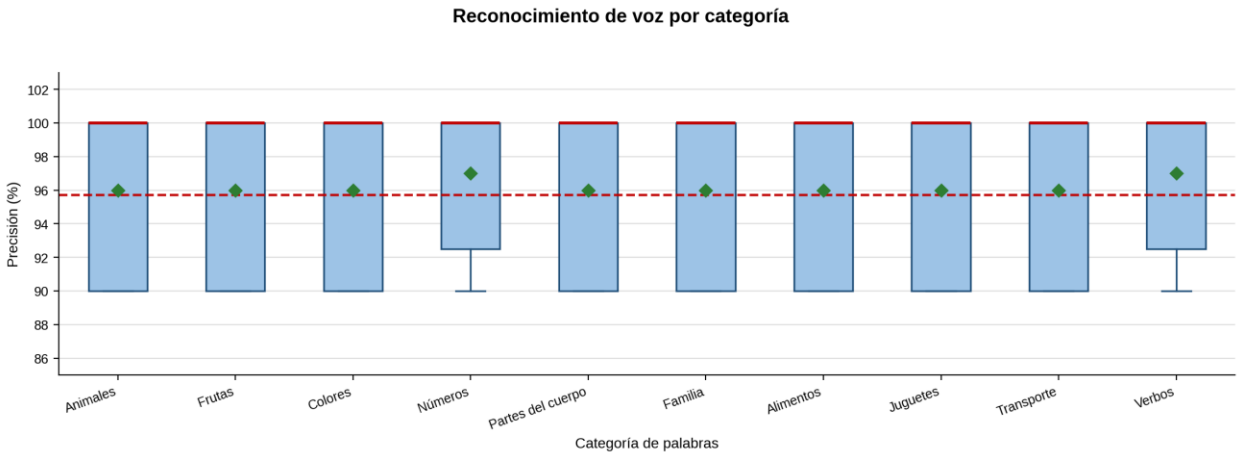
Esta sección documenta las pruebas realizadas a la interfaz gráfica del robot: el tiempo de carga de las pantallas de la HMI y el desempeño del sistema de reconocimiento de voz en tareas de pronunciación. Con ellas se comprueba la fluidez de la navegación entre pantallas y la fiabilidad del reconocimiento de voz frente a las respuestas de los niños.

Figura 61
Tiempos de carga de las pantallas del HMI



Nota. Elaboración propia.

Figura 62
Resultados de reconocimiento de voz



Nota. Elaboración propia.

Validación del prototipo final

- Total de palabras: 10
- Intentos por palabra: 10
- Total de pronunciaciones evaluadas: 1000
- Reconocimientos correctos: 957
- Precisión global del sistema: 95.7 %

Análisis de Costos

Para saber si el prototipo es viable económicamente, se hizo un análisis de los costos de los componentes y materiales usados en las etapas de diseño, construcción e implementación del robot educativo interactivo. Este análisis incluye los costos del hardware, la fabricación de la estructura, el desarrollo electrónico, la creación de la placa PCB y otros elementos necesarios para que el sistema funcione. Al estimar estos costos, se puede determinar cuánto cuesta construir el prototipo y si es factible producirlo, implementarlo y ampliar su uso en entornos educativos en el futuro.

Tabla 24
Costos de Hardware.

Componente	Cantidad	Costo Unitario ($)	Costo Total ($)
Raspberry Pi	1	199.99	199.99
Pantalla Táctil	1	99.99	99.99
Servomotor MG996R	4	13.98	55.92
Micrófono USB	1	13.99	13.99
Parlante	1	25.91	25.91
Matriz LED	2	4.20	8.40
Fuente conmutada	1	15	15
Micrófono Inalámbrico	1	13	13
Kit Mouse y Teclado Inalámbricos	1	17	17
Extensión de Corriente	1	8	8

Nota. Elaboración propia

Tabla 25
Costos Mecánicos

Elemento	Costos ($)
Filamento PLA	25
Impresiones 3D	30
Ejes y soportes	15
Acabados y pintura	10
Tornillería y accesorios	10

Nota. Elaboración propia.

Tabla 26
Costos Electrónicos de Desarrollo

Elemento	Costos ($)
Protoboards	10
Cables Jumper	5
Conectores	8
Componentes electrónicos	15

Nota. Elaboración propia.

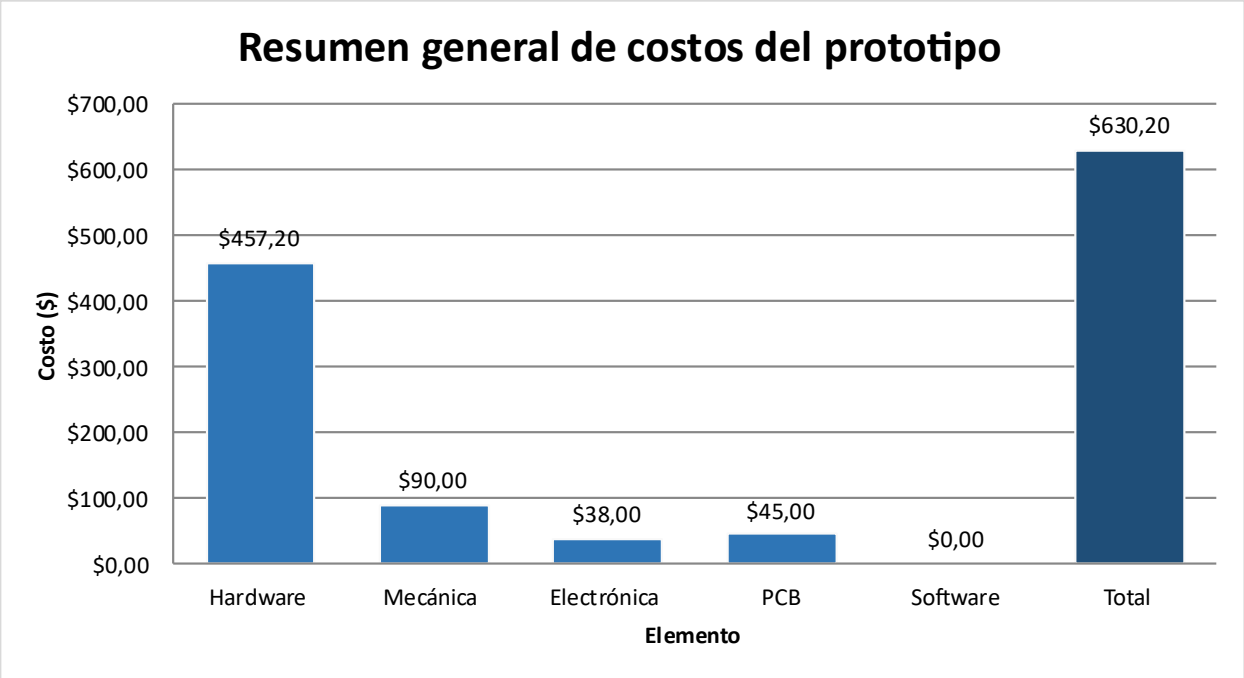
Las tecnologías utilizadas para el software fueron de uso libre y gratuitas

Tabla 27
Costos de Fabricación de PCB

Elemento	Costos ($)
Diseño	0
Fabricación	20
Componentes	15
Puntos de suelda	10

Nota. Elaboración propia.

Figura 63
Resumen general de costos



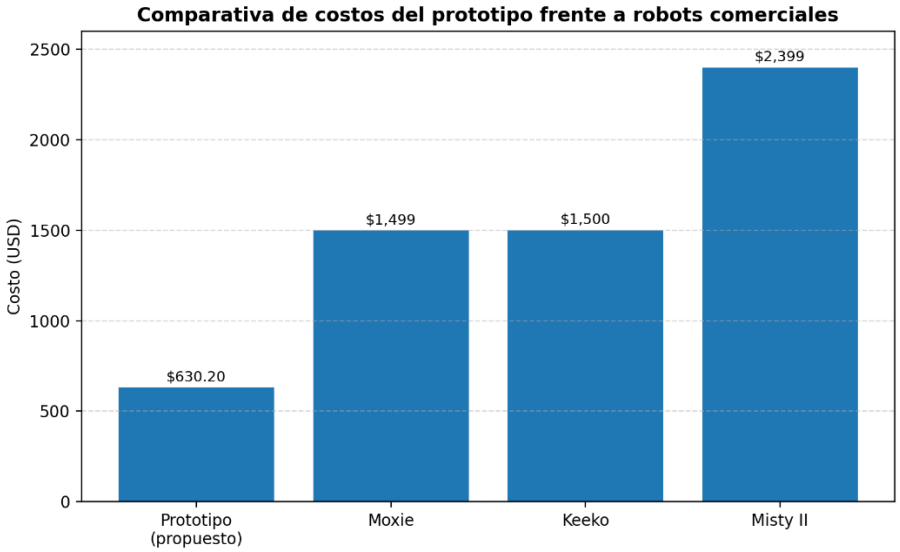
Nota. Elaboración propia.

Finalmente, si se consideran los costes en que se incurre para la realización de hardware ($457,20), fabricación mecánica ($90), desarrollo electrónico ($38), fabricación de PCB ($45) y desarrollo de software (sin coste, pues se emplean herramientas de código abierto), se puede confirmar que la cantidad necesaria para realizar el sistema del prototipo de robot educativo interactivo es de $630,20 dólares en total. Esto nos hace ver que el desarrollo del sistema es

viable rápidamente en relación con un coste al alcance del presupuesto que implica realizar un proyecto de titulación, además establece una referencia para futuras optimizaciones o para la producción futura del mismo a gran escala.

Al comparar esta cifra con robots educativos y sociales de uso real en el mercado —y no con simples juguetes electrónicos—, el prototipo resulta considerablemente más económico. Misty II, una plataforma robótica social pensada para desarrolladores, investigadores y educadores, salió al mercado en 2399 dólares (Heater, 2019). Keeko, un robot chino desplegado en más de 600 jardines infantiles, cuesta alrededor de 1500 dólares (10 000 yuanes) (Agence France-Presse, 2018). Moxie, un robot social orientado al desarrollo socioemocional infantil, se lanzó en 2020 a un precio promocional de 1499 dólares (Kara, 2020). Frente a estas cifras, el costo de 630,20 dólares del prototipo desarrollado en este trabajo resulta entre 2,4 y 3,8 veces menor, pese a integrar funciones de reconocimiento de voz, narración interactiva y corrección de pronunciación diseñadas específicamente para el español y para el rango etario de 4 a 5 años, un nicho que ninguno de esos tres robots atiende de forma nativa. Esta comparación confirma que el prototipo es técnicamente viable y, además, mucho más accesible que alternativas comerciales de propósito educativo similar. Un resumen de esta información se puede ver en la Figura 64.

Figura 64
Comparativa de precios con prototipos comerciales



Nota. Elaboración propia.

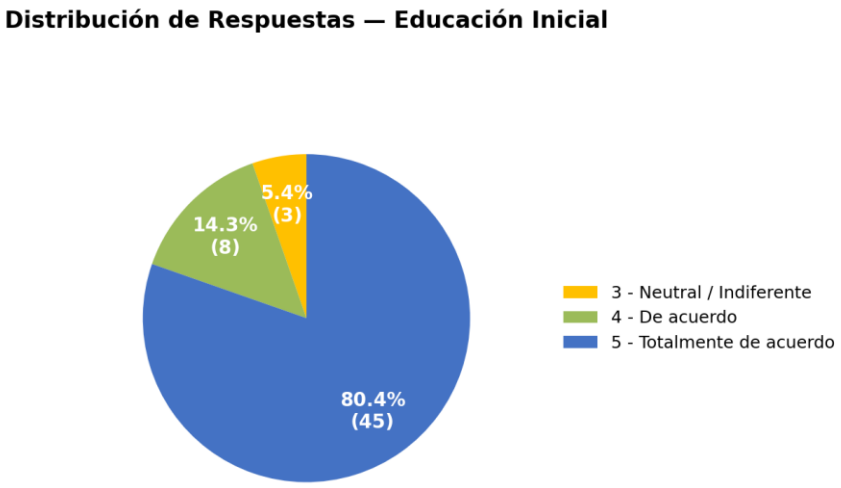
.

Validación mediante encuestas a expertos

Además de las pruebas técnicas descritas en las secciones anteriores, el prototipo se sometió a una validación cualitativa mediante encuestas de percepción, aplicadas a dos grupos de expertos con miradas complementarias sobre el proyecto. El primer grupo lo conformaron cuatro docentes de Educación Inicial, que evaluaron la pertinencia pedagógica del sistema y la experiencia de uso desde la perspectiva del aula; el segundo, cinco docentes del Departamento de Ciencias de la Energía y Mecánica, que evaluaron la robustez técnica, la integración de los subsistemas y el potencial de escalabilidad del prototipo. Entre los miembros de cada grupo se extrajeron 27 criterios de evaluación (14 de la encuesta de Educación Inicial y 13 de la encuesta de Ingeniería). Estos criterios forman parte de cada una de las dimensiones, según la escala Likert, que tiene un rango de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo).

El grupo de docentes de Educación Inicial presentó un promedio general de 4,75 sobre 5. El porcentaje que obtuvo la mayor calificación fue el de Personalización y Ampliación de Contenido (5,0); posteriormente, respecto a la dimensión de Implementación en el Aula obtuvieron la calificación de (4,88) y Retroalimentación y Motivación se resolvió en (4,75).El puntaje más bajo, aunque igualmente favorable, correspondió a la dimensión de Interfaz y Experiencia del Niño (4,5), donde las expresiones del robot —movimiento de cabeza, brazos y ojos LED— recibieron la calificación más baja de esa dimensión (4,0) sin dejar de reflejar una valoración positiva.

Figura 65
Distribución de respuestas — Educación Inicial

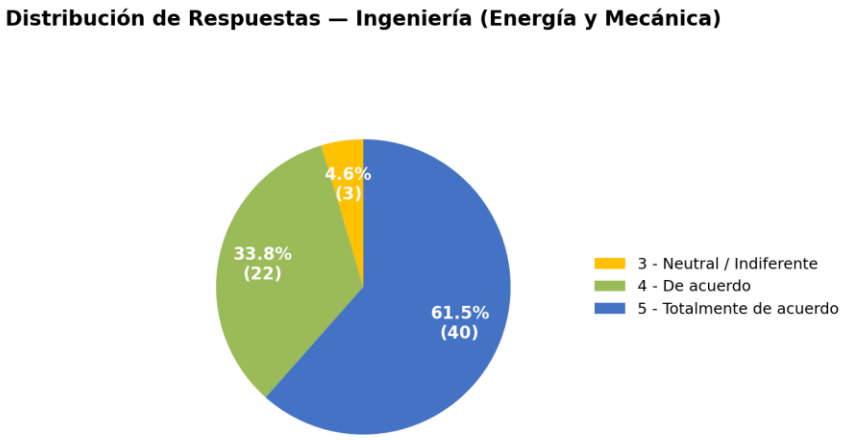


Nota. Elaboración propia.

Los docentes del Departamento de Ciencias de la Energía y Mecánica reportaron, por su parte, un promedio general de 4,57 sobre 5. Las dimensiones de Escalabilidad y Gestión de Contenido (4,8) e Integración, Robustez y Seguridad (4,7) obtuvieron las calificaciones más altas, lo que respalda la solidez estructural y el potencial de crecimiento del sistema. La dimensión de

Movimiento y Expresividad Física (4,4) fue la de menor puntaje, principalmente por la sincronización de los paneles LED de los ojos con los movimientos de cabeza y brazos, que promedió 4,0.

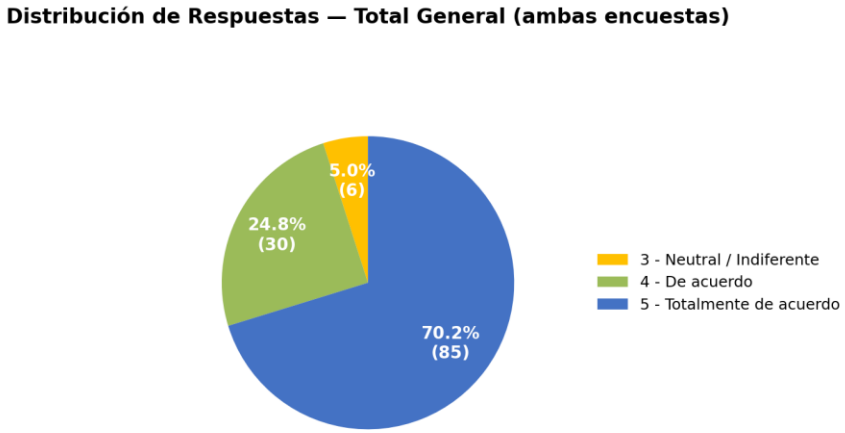
Figura 66
Distribución de respuestas — Ingeniería (Energía y Mecánica)



Nota. Elaboración propia.

Cuando se agrupan las respuestas de ambos grupos, 121 evaluaciones en total, el 70,2% se encuentra en la categoría de "Totalmente de acuerdo" (85 respuestas); el 24,8% está en la categoría de "De acuerdo" (30 respuestas), el 5,0% de las respuestas (6 de 121) es de valoración neutral; y no se encuentra valoración en la categoría de desacuerdo. Este perfil, que es coincidente entre los dos grupos evaluadores a pesar de su distinta formación, también avala tanto la adecuación pedagógica correspondiente al robot educativo como la consistente aplicación técnica.

Figura 67
Distribución de respuestas — Total General (ambas encuestas)



Nota. Elaboración propia.

En conjunto, los resultados de las encuestas complementan la validación técnica presentada en las secciones anteriores: confirman que el prototipo cumple con los parámetros de funcionamiento esperados y, además, resulta pertinente, seguro y bien recibido tanto por especialistas en pedagogía infantil como por especialistas en ingeniería mecatrónica.

Capítulo V

Conclusiones y recomendaciones

Conclusiones generales

Se realizó un estudio del arte que abarcó más de 35 fuentes académicas sobre robótica social, narración asistida por robots, reconocimiento automático del habla infantil (ASR) y retroalimentación fonética, lo que permitió identificar que, si bien existen antecedentes de robots narradores y de corrección de pronunciación por separado, no se encontraron soluciones que integren comprensión de cuentos, formación de oraciones y corrección de pronunciación en un mismo dispositivo orientado a niños de 4 a 5 años. Este vacío identificado en la literatura fundamentó el enfoque integral adoptado en el diseño del prototipo.

El sistema mecánico, electrónico y de control se diseñó y validó mediante análisis de esfuerzos y factores de seguridad. En el subsistema mecánico se aplicó un factor de seguridad mínimo de 7 en el dimensionamiento de los servomotores, obteniéndose esfuerzos máximos de 0,366 MPa en la cabeza y 0,138 MPa en el torso, muy por debajo del límite de 47 MPa del material (PLA), lo que se tradujo en un coeficiente de seguridad de 15 en el torso. En cuanto al subsistema eléctrico, se evaluó el propio factor de seguridad alcanzando un 5,51 respecto al consumo nominal (2,72 A), y un 1,03 relativo al consumo máximo estimado (14,5 A), alcanzando incluso un margen aproximado de 500 mA aun en el escenario de máxima demanda simultánea, es decir, el diseño satisface los requerimientos del propio sistema de estructuras y energética.

Se obtuvo también una interfaz gráfica (HMI) que consiste en 14 pantallas interactivas o HMI, una por cada uno de los módulos de aprendizaje (comprensión de cuentos, creación de oraciones y cumplimiento). Las pruebas de tiempo de carga, realizadas sobre 20 mediciones, arrojaron un tiempo promedio de 0,86 s (mínimo de 0,81 s y máximo de 0,91 s), lo que evidencia

una respuesta ágil adecuada para mantener la atención de niños de 4 a 5 años durante la interacción.

Los sistemas diseñados (mecánico, electrónico, de control y la interfaz gráfica) se implementaron e integraron con éxito en un prototipo funcional, fabricado mediante impresión 3D por deposición de material fundido (FDM) y ensamblaje manual. El análisis de costos determinó una inversión total de 630,20 dólares, distribuidos en hardware (457,20 USD), fabricación mecánica (90 USD), desarrollo electrónico (38 USD) y fabricación de PCB (45 USD), lo que demuestra la factibilidad técnica y económica de la implementación del robot educativo interactivo que comparado con los robots Moxie, Keeko y Misty 2, es alrededor de 2,4 veces más barato de implementar.

La validación y las pruebas de funcionamiento del prototipo confirmaron, finalmente, su fiabilidad en los distintos subsistemas evaluados. Respecto a las pruebas de expresión mecánica (felicidad y tristeza), se ha manifestado un valor de fiabilidad correspondiente al 100% (20/20 secuencias correctas) en ambas expresiones; y registrados errores angulares medios de hasta 1° y errores de porcentajes entre 0,75% y 2,5% en relación a los valores programados. Respecto a las pruebas de reconocimiento de pronunciación, se han alcanzado 957 reconocimientos correctos sobre un total de 1000 evaluaciones realizadas (100 palabras con 10 intentos cada una), que corresponde a un valor de precisión global del sistema del 95'7%. De este modo, los resultados obtenidos, sumados al tiempo de respuesta medio de 0'86 s que ofrece la correspondiente interfaz gráfica, permiten concluir que el prototipo diseñado funciona de acuerdo con lo previsto y se podría convertir en una herramienta de apoyo a la práctica de habilidades lingüísticas de niños/as de 4 a 5 años. Junto a esta validación técnica, existió una validación cualitativa a partir de encuestas de percepción administradas a cuatro docentes de Educación Inicial así como a otros cinco docentes del Departamento de Ciencias de la Energía y Mecánica,

mediante una escala de tipo Likert de 1 a 5. El primer grupo obtiene una nota media de 4,75 sobre un máximo de 5 y el segundo grupo obtiene una nota media de 4,57 sobre 5; así pues entre los 121 datos obtenidos, el 70,2% de las respuestas agruparon en la categoría "Totalmente de acuerdo" y ninguna puntuación se obtuvo en las categorías de desacuerdo. Estos resultados, que son coherentes entre los dos grupos evaluadores que tienen una distinta preparación, confirman la pertinencia didáctica del robot y su fuerte implantación técnica; además, los dos grupos evaluadores coinciden al afirmar que el Administrador de Contenido es un instrumento muy bien pensado, de hecho fue el aspecto mejor puntuado en las dos encuestas.

Recomendaciones técnicas y operativas

Conviene revisar cada cierto tiempo el apriete de los tornillos y el estado de los acoples impresos en 3D, ya que el movimiento repetido de los servomotores en brazos y cabeza tiende a aflojar las uniones con el uso. Cada vez que se desmonte o se cambie alguna pieza mecánica es importante recalibrar los ángulos de referencia de los servos, comprobando que la posición que marca el software coincida con la posición real del robot.

En el aspecto eléctrico es recomendable no realizar sesiones largas de más de 5 horas, cuando el consumo máximo sea cercano al estimado (14,5 A), ya que en ese momento el margen de seguridad de la fuente se reduce hasta un valor de 1,03; la temperatura de la fuente conmutada y de la PCB conviene vigilarla durante un uso intensivo. El robot también se comporta mejor en espacios con poco ruido ambiental: aunque el sistema filtra el zumbido eléctrico y adapta la detección del silencio al ruido ambiental del lugar, en espacios con mucho ruido el reconocimiento de voz puede seguir siendo susceptible a interferencias acústicas que disminuyan la precisión del 95,7% obtenida en las pruebas, así que es mejor no utilizarlo con otros dispositivos con micrófono activo en el espacio.

Durante las sesiones con los niños siempre debería estar presente un adulto, sea docente o cuidador, ya que el robot está pensado como apoyo al aprendizaje del lenguaje y no como reemplazo de la guía humana. Cuando no esté en uso conviene guardarlo en un lugar seco, a temperatura ambiente y lejos de la luz solar directa, para evitar que las piezas impresas en PLA se deformen. Por último, es recomendable mantener actualizados los archivos de diseño CAD, los modelos STL de impresión y el código del sistema de control, la interfaz y el reconocimiento de voz, de modo que quien retome el proyecto más adelante pueda repararlo o modificarlo sin partir de cero.

Trabajos Futuros

A partir de los resultados obtenidos y de las limitaciones identificadas durante el desarrollo de este trabajo, se plantean las siguientes líneas de investigación y mejora, con el fin de que puedan ser retomadas por futuros estudiantes o investigadores interesados en dar continuidad al proyecto.

En la parte mecánica valdría la pena sumar más grados de libertad, por ejemplo movimiento de torso o piernas, para ampliar el repertorio de gestos más allá de la felicidad y la tristeza evaluadas en este trabajo, y probar materiales alternativos o híbridos que bajen el peso del prototipo sin sacrificar el factor de seguridad estructural que se logró.

En electrónica y energía, una batería recargable tipo LiPo permitiría usar el robot de forma portátil, sin depender de un tomacorriente cercano como ocurre con la fuente conmutada actual. También sería interesante incorporar una cámara con algoritmos de visión por computador que detecten el nivel de atención o el estado emocional del niño y ajusten la interacción en tiempo real.

En cuanto a la voz, entrenar o ajustar el reconocimiento de habla con un corpus infantil en español podría superar el 95,7% de precisión obtenido en las pruebas, sobre todo para el motor local Vosk, que es el que se usa cuando no hay conexión a Internet. También convendría evaluar el desempeño del sistema en instituciones con una conexión a Internet más estable, para aprovechar con mayor frecuencia la mayor precisión que ofrece el motor en línea de Google. En cuanto a la síntesis de voz, valdría la pena explorar modelos con mayor naturalidad o con capacidad de expresar distintas emociones en el tono de la voz, si en el futuro se cuenta con hardware de mayor capacidad de procesamiento.

El contenido educativo puede mejorar su funcionalidad: más cuentos, más categorías de vocabulario, niveles de dificultad que se ajusten al desempeño de cada niño e incluso la posibilidad de extender el sistema a otros idiomas. Como paso previo al cierre del proyecto, además, será preciso completar la batería de pruebas técnicas llevadas a cabo en el contexto de precisión angular, tiempos de respuesta y reconocimiento de voz con un estudio piloto y de largo recorrido en aula real que pudiera medir la influencia que el robot ha tenido en la evolución del lenguaje por encima del resultado producido por sus subsistemas. Una aplicación para padres y para profesores que permitiera hacer un seguimiento del avance del niño y que permitiera ir actualizando de manera remota el contenido sería una buena manera de cerrar esta etapa.

Bibliografía

Agence France-Presse. (2018, August 30). *Robot teachers invade Chinese kindergartens.* Retrieved from Phys.org: https://phys.org/news/2018-08-robot-teachers-invade-chinese-kindergartens.html

Ahmad, M. I., Mubin, O., Shahid, S., & Orlando, J. (2019). Robot’s adaptive emotional feedback sustains children’s social engagement and promotes their vocabulary learning: a long-

term child–robot interaction study. *Adaptive Behavior, 27*(4), 243-266. doi:10.1177/1059712319844182

Appel, M., Lugrin, B., Kühle, M., & Heindl, C. (2021). The emotional robotic storyteller: On the influence of affect congruency on narrative transportation, robot perception, and persuasion. *Computers in Human Behavior, 120*. doi:10.1016/j.chb.2021.106749

Barry, R.-J., Neumann, M., & Neumann, D. (2025). Individual differences and young children's engagement with a social robot. Retrieved from https://www.sciencedirect.com/science/article/pii/S2949882125000234

Bertacchini, F., Demarco, F., Scuro, C., Pantano, P., & Bilotta, E. (2023). A Social Robot Connected with ChatGPT to Improve Cognitive Functioning in ASD Subjects. *Frontiers in Psychology, 14*. doi:10.3389/fpsyg.2023.1232177

Bizarro Torres, N., Luengo González, R., & Carvalho, J. L. (2018). Roamer, un robot en el aula de Educación Infantil para el desarrollo de nociones espaciales básicas / Roamer, a robot in the classroom of Early Childhood Education for the development of basic spatial notions. *RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação*, 14-28. doi:https://doi.org/10.17013/risti.28.14-28

Blavette, L., Dacunha, S., Alameda-Pineda, X., Hernández García, D., Gannot, S., Gras, F., . . . Pino, M. (2025). Acceptability and Usability of a Socially Assistive Robot Integrated With a Large Language Model for Enhanced Human-Robot Interaction in a Geriatric Care Institution: Mixed Methods Evaluation. *JMIR Human Factors, 12*(e76496). doi:10.2196/76496

Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). Social Robotics. In *Springer Handbook of Robotics* (pp. 1935-1971). doi:10.1007/978-3-319-32552-1_72

Chaparro Altamirano, D., Zavala-Yoe, R., & Ramirez-Mendoza, R. (2015). *Análisis, Diseño y Control de un Manipulador Paralelo 3-SPS-1S Utilizado en Aplicaciones de Seguridad.*

Chatzichristofis, S. A. (2023). Recent Advances in Educational Robotics. *12*(4). doi:10.3390/electronics12040925

Chen, H., Park, H., & Breazeal, C. (2020). Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Computers & Education, 150*. doi:10.1016/j.compedu.2020.103836

Chiara, V. D. (2021, Mayo 23). *Rangos de movimiento expresivos eje superior*. Retrieved from CASIOPEA: https://wiki.ead.pucv.cl/Victor_Della_Chiara_-Rangos_de_movimiento_expresivos_eje_superior.Rangos de movimiento expresivos eje superior

Conti, D., Cirasa, C., Di Nuovo, S., & Di Nuovo, A. (2020). Robot, tell me a tale! : A social robot as tool for teachers in kindergarten. *Interaction Studies, 21*(2), 220-242. doi:10.1075/is.18024.con

Esfandbod, A., Rokhi, Z., Meghdari, A. F., Taheri, A., Alemi, M., & Karimi, M. (2023). Utilización de un robot emocional capaz de sincronización labial en sesiones de terapia del habla asistida por robot para niños con trastornos del lenguaje. *International Journal of Social Robotics*, 165-183. doi:https://doi.org/10.1007/s12369-022-00946-2

Essien, G., & Parbanath, S. (2024). Exploring the World of Robot-Assisted Digital Storytelling: Trends, Models, and Educational Implications. *Journal of Information Technology Education: Research*, Art. 27.

Fang, J.-W., Ji, T.-T., Tu, Y.-F., Hwang, G.-J., Zou, D., & Chen, J. (2025). Effects of a human-robot collaborative teaching approach on preschoolers' social-emotional competence through collaboration between the child and robot. *156*. doi:10.1016/S0360-1315(25)00237-4

Feldman, H. M. (2019). The Importance of Language-Learning Environments to Child Language Outcomes. *Pediatrics, 144*(4). doi:10.1542/peds.2019-2157

Fink, J. (2012). Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In J. Fink, *Social Robotics (parte de la serie Lecture Notes in Computer Science)* (pp. 199-208). Springer. doi:10.1007/978-3-642-34103-8_20

Heater, B. (2019, January 8). *Misty's adorable robotics platform ships in April for $2,399.* Retrieved from TechCrunch: https://techcrunch.com/2019/01/08/mistys-adorable-robotics-platform-ships-in-april-for-2399/

Iio, T., Yoshikawa, Y., Ogawa, K., & Ishiguro, H. (2024). Comparison of Outcomes Between Robot-Assisted Language Learning System and Human Tutors: Focusing on Speaking Ability. *International Journal of Social Robotics, 16*, 743-761. doi:10.1007/s12369-024-01134-0

Janssens, R., Verhelst, E., Abbo, G. A., Ren, Q., Pinto Bernal, M. J., & Belpaeme, T. (2024). *Child Speech Recognition in Human-Robot Interaction: Problem Solved?* arXiv. Retrieved from https://arxiv.org/abs/2404.17394

Kara, D. (2020, August 19). *How Embodied brought Moxie to market at the lowest price possible.* Retrieved from The Robot Report: https://www.therobotreport.com/how-embodied-brought-moxie-to-market-at-the-lowest-price-possible/

Kory, J. M., & Breazeal, C. (2014). Storytelling with Robots: Learning Companions for Preschool Children's Language Development. *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 648–653). Washington, DC: IEEE. doi:10.1109/ROMAN.2014.6926325

Kory-Westlund, J. M., & Breazeal, C. (2019). A Long-Term Study of Young Children's Rapport, Social Emulation, and Language Learning With a Peer-Like Robot Playmate in Preschool. *Frontiers in Robotics and AI (abreviado "Front. Robot. AI")*, Article 81. doi:10.3389/frobt.2019.00081

Li, L., Wang, X., Wang, Q., Zhong, B., & Xu, Y. (2023). Meta–analyzing the Impacts of Social Robots for Children's Language Development: Insights from Two Decades of Research from 2003 to 2023. Retrieved from https://www.linlinliedu.com/publication/preprint/

Mazı, A., & Yıldırım, İ. O. (2025). Primary school teachers' opinions on the use of artificial intelligence in educational practices. *Social Sciences & Humanities Open, 11*. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590291125003043

Neumann, M. M. (2020). Social Robots and Young Children's Early Language and Literacy Learning. *Early Childhood Education Journal, 48*(2), 157-170. doi:10.1007/s10643-019-00997-7

Nomura, T. (2017). Robots and gender. *Gender and the Genome, 1*(1), 18-26. doi:10.1089/gg.2016.29002.nom

Papadopoulos, I., Lazzarino, R., Miah, S., Weaver, T., Thomas, B., & Koulouglioti, C. (2020). A systematic review of the literature regarding socially assistive robots in pre-tertiary education. *Computers & Education, 155*. Retrieved from https://doi.org/10.1016/j.compedu.2020.103924

Pasalidou, C., Lytridis, C., Tsinakos, A., & Fachantidis, N. (2025). Augmented Reality and Robotics in Education: A Systematic Review of the Literature on the Combination of AR and Robotics for Educational Purposes. *Smart Learning Environments, 12*. Retrieved from https://www.sciencedirect.com/science/article/pii/S2949882125000416

Rodríguez, M. (2023, Abril 29). *¿Sabes en robótica la diferencia entre grados de libertad y grados de movilidad?* Retrieved from Inesem Business School: https://www.inesem.es/revistadigital/gestion-integrada/diferencia-robotica-grados-libertad-movilidad-3

Roštšinskaja, A., Saard, M., Korts, L., Kööp, C., Kits, K., Loit, T.-L., . . . Kolk, A. (2025). Unlocking the Potential of Social Robot Pepper: A Comprehensive Evaluation of Child-Robot

Interaction. *Journal of Pediatric Health Care, 39*(4), 572-584. doi:10.1016/j.pedhc.2025.01.010

Sariyati, & Wardhani, J. D. (2023). Improving Language Skills Through Storytelling Method with Audio Visual Media in Group a Early Childhood Department. *Proceedings of the International Conference on Learning and Advanced Education (ICOLAE 2022)* (pp. 1978-1990). Atlantis Press. doi:10.2991/978-2-38476-086-2_159

Smakman, M., Vogt, P., & Konijn, E. A. (2021). “Moral considerations on social robots in education: A multi-stakeholder perspective”. *Computers & Education, 274*. doi:10.1016/j.compedu.2021.104317

Sophokleous, A., Christodoulou, P., Doitsidis, L., & Chatzichristofis, S. A. (2021). Computer Vision Meets Educational Robotics. *Electronics*, 730. doi:https://doi.org/10.3390/electronics10060730

Spaulding, S., Chen, H., Ali, S., Kulinski, M., & Breazeal, C. (2018). A Social Robot System for Modeling Children’s Word Pronunciation: Socially Interactive Agents Track. *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)* (pp. 1658-1666). Stockholm: IFAAMAS.

Wang, X., Wang, Y., Han, Z., Wang, Y., Duan, Z., Zhang, Z., & … Alsudais, T. A. (2025). Social Robots for Child Development: Research Hotspots, Topic Modeling, and Collaborations. *Humanities and Social Sciences Communications, 12*(Art 1411). doi:https://doi.org/10.1038/s41599-025-05752-5

Wang, Z., Law, T. S.-T., & Yeung, S. S. (2025). Let robots tell stories: Using social robots as storytellers to promote language learning among young children. *Computers in Human Behavior: Artificial Humans*. doi:10.1016/j.chbah.2025.100210

Zhang, N., Xu, J., Zhang, X., Wang, Y., & al., e. (2024). Social Robots Supporting Children's Learning and Development: Bibliometric and Visual Analysis. *Education and Information Technologies, 29*, 12,115–12,142. doi:10.1007/s10639-023-12362-8

Zorrilla-Puerto, J., Lores-Gómez, B., Martínez-Requejo, S., & Ruiz-Lázaro, J. (2023). El papel de la robótica en Educación Infantil: revisión sistemática para el desarrollo de habilidades. *RiiTE. Revista Interuniversitaria de Investigación en Tecnología Educativa*(15), 188-194. doi:10.6018/riite.586601

Apéndices