

**DESARROLLO DE UN ASISTENTE VIRTUAL INTELIGENTE PARA APOYO AL
PROCESO DE APRENDIZAJE DE LA ASIGNATURA DE INGLÉS NIVEL I PARA
LOS ESTUDIANTES DEL PROGRAMA DE SISTEMAS DE LA UNIVERSIDAD DE
PAMPLONA AMPLIACIÓN VILLA DEL ROSARIO**

YONATHAN JAVIER JAIMES MANTILLA



**UNIVERSIDAD DE PAMPLONA
FACULTAD DE INGENIERÍAS Y ARQUITECTURA
DEPARTAMENTO DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA, SISTEMAS Y
TELECOMUNICACIONES
INGENIERÍA DE SISTEMAS
VILLA DEL ROSARIO**

2022

**DESARROLLO DE UN ASISTENTE VIRTUAL INTELIGENTE PARA APOYO AL
PROCESO DE APRENDIZAJE DE LA ASIGNATURA DE INGLÉS NIVEL I DE LOS
ESTUDIANTES DEL PROGRAMA DE SISTEMAS DE LA UNIVERSIDAD DE
PAMPLONA AMPLIACIÓN VILLA DEL ROSARIO**

YONATHAN JAVIER JAIMES MANTILLA

**PROYECTO DE GRADO PARA OPTAR POR EL TÍTULO DE INGENIERO DE
SISTEMAS**

DIRECTOR

RICHARD ELISEO MENDOZA GÁFARO

UNIVERSIDAD DE PAMPLONA

FACULTAD DE INGENIERÍAS Y ARQUITECTURA

**DEPARTAMENTO DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA, SISTEMAS Y
TELECOMUNICACIONES**

INGENIERÍA DE SISTEMAS

VILLA DEL ROSARIO

2022

AUTOR

YONATHAN JAVIER JAIMES MANTILLA

RICHARD ELISEO MENDOZA GÁFARO

INGENIERO DE SISTEMAS

INGENIERÍA DE SISTEMAS

**DEPARTAMENTO DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA, SISTEMAS Y
TELECOMUNICACIONES**

FACULTAD DE INGENIERÍA Y ARQUITECTURA



UNIVERSIDAD DE PAMPLONA

VILLA DEL ROSARIO, 2022

Dedicatoria

A Dios quiero dar gracias por la oportunidad que me dio de estudiar, por la fortaleza y ayuda en los momentos más difíciles de mi desarrollo académico, y por su Santa guianza.

También quiero agradecer a mis queridos padres Víctor Manuel Jaimes Rangel y Yolima Mantilla Montagut, quienes desde pequeño me han apoyado en mis estudios, pues gracias a Dios primeramente y a su sacrificio me han podido dar las herramientas necesarias para aprender. Doy gracias a Dios por ellos.

Quiero agradecer también a todos aquellos docentes que me han acompañado a lo largo de la carrera, que el Señor les bendiga, y que sigan realizando una buena función como lo es la enseñanza.

Resumen

El desarrollo de este proyecto tiene como finalidad dar apoyo con el uso de un asistente virtual inteligente a estudiantes de Ingeniería de Sistemas de la Universidad de Pamplona, que, por dificultad del mismo inglés, se les complica el habla, pues el idioma al tener unos fonemas y estructura diferente cuando se habla, fuera como si se estuviera aprendiendo a hablar de nuevo, pero esta vez en un idioma distinto.

Añadido a lo anterior se presenta la barrera del miedo, que impide que el aprendizaje total del idioma se lleve a cabo, pues este factor hace difícil para la persona que intenta comunicarse, el poder expresarse así mismo con otros que hablan inglés.

Como aporte a la solución de este problema, se idea el uso de un asistente virtual inteligente con el fin de emular las interacciones conversacionales entre estudiante y hablante natural del inglés. Este asistente virtual inteligente es el de IBM Watson, también se utilizará la plataforma Unity, la aplicación de desarrollo de bases de datos SQLite; con las cuales se desarrollará un asistente virtual inteligente enfocado el desarrollo de la conversación del idioma inglés para estudiantes. Este proceso está pensado en un método simple donde la persona interactúa, escucha y responde dónde la máquina toma lo que la persona ha dicho y busca la intención dentro de la frase, por lo cual hay unas intenciones predefinidas en el asistente y de acuerdo con las intenciones que el asistente detecte, va a dar una respuesta a lo que está diciendo el usuario.

Este proyecto se realizará con una metodología híbrida, la cual es una combinación de algunos elementos de la metodología Scrum y la metodología Kanban.

Palabras claves: Asistente virtual, Speech to text, Aprendizaje, Inglés, Text to speech.

Abstract

The purpose of the development of this project is to provide help to the students of System Engineering of the University of Pamplona using an intelligent virtual assistant, due of difficulties of the same English, it is not easy for them to speak it, since the language has different phonemes and structure when it is spoken, it would be like if someone were learning to speak again, but this time in a different language.

Added to the above, there is the wall of fear, which stops the full learning of the language, because of this factor makes it difficult for the person who tries to communicate, to be able to express himself/herself with others who speak English.

As a contribution to the solution of this problem, the idea of using an intelligent virtual assistant is conceived in order to emulate the conversational interactions between students and natural English speakers. This intelligent virtual assistant is IBM Watson, also the Unity platform, the database development application SQLite will be used; with which an intelligent virtual assistant will be developed with focus on the improvement of the English language conversation for students. This process is minded in a simple method where the person interacts, listens and answers, then the machine takes what the person has said and looks for the intention within the sentence, for which there are predefined intentions in the assistant and according to the intentions that the assistant detects, it will give a response to what the user is saying.

This project will be done with a hybrid methodology, which is a combination of some elements of the Scrum methodology and the Kanban methodology.

Keywords: Virtual assistant, Speech to text, Learning, English, Text to speech.

TABLA DE CONTENIDO

Introducción	15
Capítulo 1: Generalidades.....	16
Planteamiento Del Problema	16
Formulación del problema.....	16
Descripción del problema.....	16
Justificación	19
Delimitación	20
Objetivo General.....	20
Objetivos Específicos	20
Acotaciones	20
Marco Referencial	21
Antecedentes.....	22
Estado Del Arte	29
Marco Teórico	31
Metodología De Desarrollo.....	35
Scrum:.....	35
Kanban:.....	36
Metodología Híbrida:	36
Implementación De La Metodología De Desarrollo	38
Capítulo 2: Análisis de tecnologías para desarrollo de asistentes inteligentes	40
Tecnologías para diseño de aplicaciones con asistentes virtuales.....	40
Unity	40
C#.....	41
SQLite.....	42
Watson Assistant, Speech to Text y Text to Speech	42
Capítulo 3: Definición de intenciones enfocadas en el desarrollo conversacional	43
Intenciones	43

Capítulo 4: Desarrollo de la aplicación.....	48
Asistente	48
Descripción Del Proceso De Comunicación Con Watson Assistant.....	48
Diagramas UML	51
Identificación de actividades de cada actor	57
Tabla De Descripción De Cada Caso De Uso	59
Aplicación	67
Funcionamiento	68
Código	74
IBM Watson Assistnat	77
Intenciones:.....	77
Entidades:	80
Diálogos:.....	82
Conexión.....	90
Base de datos	91
Capítulo 5: Validación	94
Testeo y Datos Recolectados.....	94
Conclusiones	106
Recomendaciones	108
Glosario.....	109
Referencias.....	111
ANEXOS	119

Listado de Diagramas

Diagrama 1: Diagrama de actividades para la práctica de Conversación	51
Diagrama 2 Diagrama de estado de un usuario entrado a practicar con el asistente	52
Diagrama 3 Diagrama general de casos de uso de la Aplicación	53
Diagrama 4 Diagrama de casos de uso Usuario.....	54
Diagrama 5 Diagrama de casos de uso Watson	55
Diagrama 6 Diagrama de casos de uso Speech To Text.....	55
Diagrama 7 Diagrama de casos de uso Text to Speech	56
Diagrama 8 Diagrama de casos de uso Base de Datos	56
Diagrama 9 Diagrama de secuencia del usuario	65
Diagrama 10 Diagrama de despliegue	66

Listado de Tablas

Tabla 1 Actividades del usuario.....	57
Tabla 2 Actividades de IBM Watson Assistant	57
Tabla 3 Actividades de Text to Speech.....	58
Tabla 4 Actividades de Speech to Text.....	58
Tabla 5 Actividades de Base de datos.....	58
Tabla 6 Descripción del caso de uso del usuario	59
Tabla 7 Descripción del caso de uso de Asistente Watson	60
Tabla 8 Descripción del caso de uso del servicio Text to Speech	61
Tabla 9 Descripción del caso de uso del servicio Speech to Text	62
Tabla 10 Descripción del caso de uso de la Base de Datos	63
Tabla 11 Intenciones utilizadas para el asistente	78
Tabla 12 Colores guía	78
Tabla 13 Entidades utilizadas por el asistente	81

Listado de Ilustraciones

Ilustración 1 Languages de programación usuados en la inteligencia artificial	40
Ilustración 2 Unity herramienta de desarrollo de video juegos	41
Ilustración 3 Csharp	41
Ilustración 4 Gestor de base de datos SQLite	42
Ilustración 5 Servicios de Watson Assistant, Speech to Text y Text to Speech.....	42
Ilustración 6 Flujo de conversación con el asistente.....	49
Ilustración 7 Conexión de la aplicación con los componentes que utiliza	50
Ilustración 8 Inicio de sesión	68
Ilustración 9 Registro.....	69
Ilustración 10 Menú	70
Ilustración 11 Ayuda.....	71
Ilustración 12 Configuraciones	71
Ilustración 13 Puntajes	72
Ilustración 14 Nivel 1.....	72
Ilustración 15 Nivel 2.....	73
Ilustración 16 Nivel 3.....	73
Ilustración 17 Credenciales para conectar con IBM Watson.....	74
Ilustración 18 Credenciales del servicio Speech to Text	74
Ilustración 19 Credenciales del servicio Text to Speech	74
Ilustración 20 Creación del servicio TTS	75
Ilustración 21 Servicio STT	75
Ilustración 22 Creación del servicio del asistente IBM Watson	76
Ilustración 23 Intenciones en la plataforma IBM Cloud.....	77
Ilustración 24 Entidades en la plataforma IBM Cloud	80
Ilustración 25 Diálogos en la plataforma IBM Cloud.....	82
Ilustración 26 Nodo Welcome	83
Ilustración 27 Variedad de saludos en el nodo Welcome	83
Ilustración 28 Nodo How are thing.....	84

Ilustración 29 Condiciones y respuestas del nodo How are thing	84
Ilustración 30 Nodo About What Do You Want to Talk	85
Ilustración 31 Condiciones y respuestas del nodo About What Do You Want to Talk.....	85
Ilustración 32 Condiciones y respuestas del nodo Conversation.....	86
Ilustración 33 Condiciones y respuestas del nodo About football.....	87
Ilustración 34 Condiciones y respuestas del nodo Like football	88
Ilustración 35 Nodo End_Conversation.....	89
Ilustración 36 Dos tablas mostradas en DB Browser for SQLite	91
Ilustración 37 Base de datos relacional.....	92
Ilustración 38 Interacción entre la base de datos y la aplicación	92

Listado de Graficas

Grafica 1 ¿A qué programa pertenece?.....	94
Grafica 2 ¿Considera usted que el aprendizaje del idioma inglés es importante para su vida profesional?.....	95
Grafica 3 ¿La aplicación fue intuitiva?	96
Grafica 4 ¿Es la aplicación usable?	97
Grafica 5 ¿Considera que la aplicación fortalece el aprendizaje del idioma inglés en la carrera?98	
Grafica 6 ¿Hubo botón para cerrar la aplicación?	99
Grafica 7 ¿Fue la información entendible?	100
Grafica 8 ¿Tiene la aplicación manual de instrucciones?	101
Grafica 9 ¿Fue legible el texto en la aplicación?	102
Grafica 10 ¿Considera que la fuente de texto utilizada es adecuada para la aplicación?	103
Grafica 11 ¿Considera que la paleta de colores usada es adecuada para la aplicación?.....	104
Grafica 12 ¿Qué porcentaje entendió el asistente de lo que usted habló?	105

Introducción

El aprendizaje de un idioma no es fácil, comenzando desde su estructura y forma, pues un idioma es una forma distinta de decir lo mismo, es otra manera de expresar una idea. Para un aprendiz de un idioma el objetivo es entenderlo y hablarlo, pero en su proceso se encuentra con muchos obstáculos, uno de esos es el habla el cual es realmente importante en el idioma, porque a través de este se expresan ideas, sentimientos y demás, pero la dificultad de poder desarrollar el habla radica en la falta de oportunidades de interacciones humanas, normalmente cuando se está aprendiendo ocurren ciertas interacciones con los docentes, pero estas se quedan cortas por la disponibilidad de tiempo que existe.

Entendiendo la importancia de las interacciones comunicativas en el desarrollo del habla en estudiantes de inglés, se hace necesario buscar alternativas que permitan a los estudiantes desarrollar su comunicación en el idioma objetivo. Por tal motivo se plantea el desarrollo de un asistente virtual inteligente con la tecnología del asistente de IBM Watson, que permita realizar estas interacciones, donde el estudiante pueda hablar y recibir respuestas, de la misma manera practicar la pronunciación y desarrollar el habla.

Para este tipo de desarrollo se utilizar tecnologías de reconocimiento del lenguaje humano como los son Text to Speak y Speak to text, los cuales mediante unos procesos matemáticos permiten que las ondas de sonido se transformen en señales digitales que la máquina pueda entender y analizar, de donde el asistente analizará el texto y de acuerdo con las intenciones dará una respuesta al usuario.

Capítulo 1: Generalidades

Planteamiento Del Problema

Formulación del problema.

¿Cómo se puede mejorar el proceso de aprendizaje del idioma inglés en la asignatura de inglés básico I utilizando un asistente virtual con los estudiantes de Ingeniería de Sistemas de la Universidad de Pamplona ampliación Villa del Rosario?

Descripción del problema.

En esta parte, se tiene que el inglés es uno de los idiomas más utilizados en el mundo para hacer negocios, ya que está entre los cinco primeros lenguajes con ese propósito (Quintanilla, 2020). Hay muchos países alrededor del mundo, y algunos tienen su propia lengua, y en ocasiones esa lengua tiene distintos dialectos, presentándose un total de más de 7000 lenguas y dialectos hablados en el mundo (Klappenbach, 2022) una persona con la intención de extenderse profesionalmente necesitaría aprender muchas lenguas, o buscar una que sea entendida por la mayoría.

A partir de esto se entiende que aprender inglés es importante, porque ayuda a la transmisión del mensaje por ser una lengua bastante esparcida y como segunda lengua adquirida es esencial para aquellos cuya primera lengua es distinta a la hablada en la comunidad, porque tiene un papel comunicativo en esa sociedad (Ellis et al., 1994). Pero hay un tópico en la adquisición o aprendizaje del inglés que presenta algunas complicaciones y es a la hora de hacerlo verbal.

Cuando una persona habla con otra, está expresando ideas, sentimientos, y eso se hace a través de fonemas que son los sonidos que componen las palabras que decimos. La dificultad se presenta al intentar hablar en un idioma que no es el nativo, porque existen fonemas que pueden ser desconocidos, que son difíciles, o que son raros porque hay fonemas que son muy extraños. Sumando, la capacidad de producir frases para comunicar puede llegar a ser baja, debido a la falta de práctica. Esto hace que algunas personas tengan dificultades cuando van a expresarse, y la desconfianza, el miedo y la ansiedad aparecen porque no tienen la práctica necesaria, según un estudio hecho en la Universidad de Kazakastan, la ansiedad es uno de los factores que juega en contra del aprendizaje de idiomas (Suleimenova, 2013).

También está la posibilidad de que no tengan el tiempo para realizar este tipo de encuentros, esto puede resultar en un estancamiento, ellos solamente entienden el idioma, son capaces de leer, escribir y entender lo que escuchan, pero en la parte del habla no la desarrollan bien, lo cual puede llevar a desanimar a los estudiantes, porque no se sienten seguros o capaces de poder expresarse en el inglés.

Otra situación que se presenta con la comunicación en inglés, y en casi cualquier idioma que se está aprendiendo, es que cuando se está hablando, el estudiante del inglés espera que la otra persona (nativo del idioma) corrija sus errores, cosa que casi nunca sucede, pues ellos no están tan interesados en corregir cada error que el estudiante haga, pues el enfoque de ellos es transmitir el mensaje. Son pocos los casos donde se pueda encontrar, alguien interesado en ayudar con la corrección del inglés, a menos que se pague por eso.

Lo anterior soporta que el idioma inglés es importante para el desarrollo profesional de las diferentes carreras, y principalmente Ingeniería de Sistemas, ya que los perfiles más solicitados son los que se complementan con la exigencia de la certificación B1 y B2 de inglés en el campo laboral. El proyecto pretende contribuir al proceso de aprendizaje mediante el desarrollo de un asistente virtual como apoyo a los estudiantes de Ingeniería de Sistemas de la asignatura nivel Inglés básico I de la Universidad de Pamplona ampliación Villa del Rosario. Con el desarrollo de esta herramienta se mejorará la motivación de los estudiantes haciendo la cátedra más significativa aplicando intenciones mediante las cuales los estudiantes fortalecerán el inglés conversacional.

Justificación

El aprendizaje de un idioma como el inglés es importante, y más si es utilizado con frecuencia en muchas partes del mundo. En la formación profesional es imperativo tener competencias en este idioma, es decir ser capaz de escucharlo, entenderlo y hablarlo, pero para lograr esto en el proceso hay obstáculos que deben ser superados.

En los últimos años se han estado utilizando las nuevas tecnologías para dar soporte al aprendizaje de idiomas, y ayudar a superar algunos de esos obstáculos, pero son muy pocas las tecnologías que ofrecen una asistencia a nivel conversacional. Hay muchas plataformas de aprendizaje como Babbel, Memrise, Drops, Duolingo, Lingo, Say hello, Cake, BBC Learning English, ABA, LearnMatch, EWA, entre otras, que ofrecen el aprendizaje de idiomas, y que ayudan al mejoramiento del inglés en los estudiantes, pero yendo más allá es no solo aprender un idioma, sino ser capaz de transmitir información.

Por lo tanto, surge la necesidad del desarrollo de un asistente virtual inteligente que tenga como objetivo ayudar al proceso del aprendizaje de los estudiantes de Ingeniería de Sistemas que cursan la asignatura Inglés básico I, ya que este ayuda al entrenamiento de la comunicación en el idioma inglés, sin tomar mucho tiempo de la persona, siendo utilizado mientras el sujeto está en cualquier sitio siempre y cuando cuente con los elementos adecuados para realizar la conexión con el asistente.

Delimitación

Objetivo General

Desarrollar un Asistente Virtual Inteligente para apoyo al proceso de aprendizaje de la asignatura de Inglés nivel I de los estudiantes del programa de Ingeniería de Sistemas de la Universidad de Pamplona ampliación Villa del Rosario.

Objetivos Específicos

1. Analizar las tecnologías para diseño, desarrollo de un asistente virtual inteligente.
2. Definir un conjunto de intenciones enfocado al dominio conversacional del idioma Inglés para sostener una conversación con un usuario del asistente virtual.
3. Desarrollar un asistente virtual empleando tecnologías de speech to text, IBM Watson Assistant y text to speech.
4. Validar la implementación del asistente virtual mediante pruebas con estudiantes de la asignatura Inglés nivel I del programa de Ingeniería de Sistemas.

Acotaciones

Los aspectos que limitan este proyecto, debido a tiempo, presupuesto y complejidad son:

1. El proyecto se encuentra restringido a un modelo básico para Windows 10 hecho con tecnología IBM Watson.
2. Las prácticas se realizarán con estudiantes avanzados del inglés de nivel I Y III de Ingeniería de Sistemas de la Universidad de Pamplona extensión Villa del Rosario
3. Las tecnologías para el desarrollo Unity 3D, Watson Assistant, SQLite.

Marco Referencial

La inteligencia artificial (AI) es una rama de la ciencia y tecnología que desarrolla máquinas las cuales pueden realizar actividades donde es requerida parte de la inteligencia natural poseída por humanos (Su-Hie et al., 2022). AI se describe como la capacidad de análisis de cualquier sistema que comprende un entorno dado y la actuación de forma inteligente para maximizar las posibilidades de alcanzar los objetivos propuestos.

Estas máquinas intentan imitar las habilidades cognitivas de los seres humanos, como el aprendizaje y la resolución de problemas.

Algunos ejemplos de AI son los motores de búsqueda web avanzados tales como Bing, también los sistemas de recomendación de contenido como lo son YouTube y Netflix, y los asistentes virtuales inteligentes como Siri, Watson Assistant, Cortana y Alexa, donde se añade a esto lo automóviles autónomos como es el caso de Tesla, donde se puede ver la toma de decisiones inteligente, estos son solo algunos ejemplos de aplicaciones de AI.

Dentro de estas tecnologías existen aquellas que buscan entender el lenguaje humano, desde hace décadas los seres humanos han trabajado en sistemas que puedan reconocer los y comandos dados por voz, los cuales se han venido mejorando, al punto de que estos sistemas de reconocimiento de las ondas de la voz han permitido que se puedan transformar un audio a texto digital; esto se ha utilizado como una forma de escritura, aunque este sistema todavía tiene algunos inconvenientes, pues el habla humana es compleja. Añadido a esto los sistemas autónomos inteligentes, que reconocen intenciones en el texto, permiten que una persona con el habla de órdenes a una máquina.

Estas intenciones en una frase pueden ser detectadas por una máquina y así dar una respuesta acorde con lo que el usuario esté pidiendo, un ejemplo de estos son los chatbots los cuales son chat robots, que responden acorde a una serie de intenciones y entidades, que de acuerdo a los diálogos antes creados dan respuesta al usuario.

Estas máquinas se combinan con un concepto de la inteligencia artificial que se conoce como text to speech que se puede traducir como de texto a habla, los cuales unidos pueden conformar un asistente inteligente que responde casi como un humano, y que puede llegar establecer conversaciones sencillas.

Antecedentes

En más o menos 10 años, han venido apareciendo asistentes virtuales inteligentes, entre estos está Google Assistant, Cortana, Siri, Alexa y Watson Assistant que ayudan a los usuarios a realizar tareas básicas en los computadores, celulares o dispositivos electrónicos, y también pueden realizar algunas tareas de gestión. En el año 2006 Watson fue creado, pero inicialmente era un Chatbot con la finalidad de ganar un juego de preguntas y respuestas, desde ahí fue evolucionado (B, 2020), después en el 2011 Apple presentó su asistente virtual, que se conoce como Siri, dos años después, en el 2013 Microsoft presentó a Cortana, pero se lanzó en el 2015. En el año 2012 Google presentó Google Now, que era un asistente para hacer búsquedas por voz, pero no bidireccionales, no fue hasta mayo del 2016 que fue presentado Google Assistant. Estos no son los únicos que han trabajado con asistentes virtuales, también Amazon ha venido desarrollando Alexa, que fue presentada en el 2015, y Telefónica que ha lanzado su propio asistente con el fin de mejorar su relación con los clientes en algunos países de Europa y Latino América en el año 2018 (Cruz, 2019).

Pero mucho antes de estos había un sistema que se remonta al menos medio siglo, Audrey que era una tecnología capaz de reconocer números absolutos y el habla, fue desarrollado por Laboratorios Bell en 1952. Una década más tarde IBM desarrolló Shoebox que procesaba operaciones matemáticas simples, como suma y resta de números. Tiempo después apareció Harpy que entendía al menos 1000 palabras. En los 80 aparecieron los sistemas Hidden Markov Model los cuales se enfocan en el análisis de frases. Hacia los 90 apareció el reconocimiento de voz en el ordenador y 20 años más tarde, apareció el primer asistente virtual digital avanzado, Siri (Alisys, 2017).

Siri:

Es un asistente virtual que trabaja con tecnologías como, reconocimiento de habla y el procesamiento natural de la lengua. Por este tipo de herramientas se toma el audio y se pasa a una forma textual (Hoy, 2018), la intención de Siri es ayudar al usuario a tener cosas hechas, por ejemplo, el envío de mensajes o emails, la reservación en un restaurante entre otras más. Todo esto ordenado por voz y con manos libres. Esta aplicación tiene acceso a las demás aplicaciones utilizadas por el usuario, con el fin de completar todas las actividades que le son ordenadas realizar (O'Boyle, 2021).

Siri usa tecnologías de inteligencia artificial como el Procesamiento de Lenguaje Natural y reconocimiento del habla, agregado a esto, también cuenta con grandes colecciones de muestras de voz de muchas personas, que permiten entrenar la, haciendo le capaz de reconocer una considerable variedad de acentos y voz humana (Dudkin, 2018).

Google Assistant:

Google Assistant es el asistente de Google. Cuando fue lanzado, hacía parte de Google Now, y estaba diseñado para ser personal. Inicialmente, Google Now brindó información que fue relevante. Conocía el lugar de trabajo del usuario, también las reuniones y viajes, y no solo eso, sino que conocía equipos deportivos de preferencia para poder presentarte información que importara (Tillman, 2021).

El asistente de Google es uno de los asistentes más utilizados en el mundo, con él las personas pueden agendar reunión, crear recordatorios, hacer búsquedas por voz y también escritas, lo que significa que el asistente puede leer texto directamente escrito por el usuario o también recibir comando de voz. Como su compañero anterior Siri, Google Assistant utiliza el Procesamiento de lenguaje natural, que usa lingüística computacional, una combinación entre la ciencia que estudia como los idiomas funcionan y las ciencias de la computación (Fawzy, 2020).

Cortana:

“Cortana es el asistente de productividad personal de Microsoft que le ayuda a ahorrar tiempo y centrar la atención en lo que más importa” (Microsoft, n.d.).

Cortana es uno de los asistentes más comunes en el mercado, al igual que Siri y Google Assistant, realiza tareas por comando de voz, como abrir aplicaciones, documentos entre algunas otras actividades que puede realizar. La tecnología de esta asistente virtual incluye “reconocimiento del habla, Procesamiento del Lenguaje Natural, modelado de dialogo entre hombre maquina y generación de lenguaje hablado” de acuerdo con Chansanchai (2014) Microsoft New Center Staff.

Alexa:

Puede realizar tareas, ejecutar comandos que alguien da, también puede dar entretenimiento, información, y ayuda al usuario con sus capacidades de inteligencia artificial (Microsoft, n.d.). Alexa también funciona con NLP (Procesamiento del Lenguaje Natural) cuando un usuario da una orden, el audio de la orden es capturado por un micrófono y enviado a Servicio de voz de Alexa, el cual convertirá el audio grabado en comandos, que de acuerdo a la petición del usuario le dará respuesta. Cuando un usuario da un comando entra en el escenario la “palabra clave”, la cual permite que el asistente sepa que está siendo invocado, después entra en juego el procesamiento de la señal, cuya función es eliminar el sonido de fondo, para obtener un resultado más limpio, siguiente el NLP, que permite pasar el audio a texto. Después el Entendimiento del Lenguaje Natural entra con un papel de reconocimiento de intenciones, convirtiendo a representaciones con sentido el texto que ha recibido. Una vez el asistente encuentra el sentido en el comando, realiza la tarea y da una respuesta de confirmación al usuario usando Generación del Lenguaje Natural, que le permite dar expresiones con sentido en el idioma utilizado (Agarwal, 2021).

IBM Watson Assistant:

Está desarrollado para ser utilizado por entidades de diferentes tipos como las prestadoras de servicios, pues se enfoca en la velocidad de la asistencia al cliente, y también brinda asistencia al operador, utilizando para realizar estas operaciones, inteligencia artificial y procesamiento de lenguaje natural (IBM, n.d.). De acuerdo con su sitio oficial, publicaciones de estudios sobre aprendizaje de máquinas muestra que IBM Watson en comparación con otras soluciones que le son competitivas ha sido un 14.7 % más preciso.

Como sus homólogos Watson Assistant usa modelos NLP para evaluar la experiencia con los usuarios. También ha integrado un modelo mejorado de detección de intenciones, que está siendo ofrecido como versión beta con diálogos en inglés, el cual es más rápido y preciso. Este nuevo modelo contiene machine learning, transfer learning y técnicas de deep learning en un modelo cohesivo que es altamente eficiente en brindar una respuesta rápida (Corporación IBM, 2022).

Speech to text

La tecnología de habla a texto no es nueva, sino que desde la década de los 50's ha estado en los laboratorios de algunas empresas como Bell. Desde entonces se han venido desarrollando hasta volverse más precisa.

La tecnología Speech to text se basa en el reconocimiento de las palabras transmitidas en audio, las cuales son procesadas con algoritmos que las analizan con otras ya pregrabadas, buscando equivalencias, estas equivalencias son tenidas en cuenta cuando se van a elegir los caracteres que va a representar el audio o la voz.

Pero antes que la voz humana sea finalmente pasada a texto, ocurren una serie de procesos donde se determina cuáles son las posibles palabras que están siendo utilizadas. La mayoría de estos softwares usan Los Modelos Escondidos de Markov, estos ayudan a estimar con probabilidades cuáles son las palabras en el audio, añadido a eso se utilizan modelos de redes neuronales los cuales trabajan bien con la clasificación de fonemas, reconocimiento de palabras aislado y adaptación al hablante (Trivedi et al., n.d.).

Estos softwares que toman audio y lo dividen en secciones, analizando cada uno de los segmentos con ayuda de algoritmos que tratan de determinar cuál es la letra o la palabra que

más se asemeja por probabilidad a la del audio, y transcribiendo esos sonidos en texto (Lyden, 2017).

Text to speech:

Desde tiempos antiguos la tecnología de texto a habla fue concebida y se hicieron algunos inventos, esta idea estuvo en la mente muchas personas, los cuales trabajaron en máquinas que pudieran convertir un texto a un sonido humano.

Se tienen ejemplos de hace unos cientos de años atrás, como las cabezas hablantes, unos dispositivos antiguos, que fueron hechas por Gerbert, Roger Bacon y Albertus Magnus hace unos 900 años atrás por espacio de más de 100 años. En los años 1700 el científico danés Christian Kressenstein trabajó en la construcción de un modelo de voz humana en la academia rusa de ciencias que podía pronunciar las cinco vocales. Esto fue seguido por una serie de máquinas que hacían sonidos como la de Wolfgang que podía producir sonidos de consonantes y también sonidos vocálicos. En el año 1837 Charles Wheatstone desarrolló una máquina hablante basada en el diseño de la máquina de Wolfgang.

En los años 30 laboratorios Bell construyó el VOCODER, un teclado operado electrónicamente que analizaba el habla y sintetizaba lo que era dicho, fue exhibido en el año 1939 en New York World's fair de acuerdo con (Gold et al., 2011).

Los sintetizadores de habla al principio sonaron robóticos, pero últimamente han mejorado sonando más natural, en donde determinar si es un humano o un robot se ha hecho cada vez más difícil. Estos sintetizadores llegaron a los computadores en la década de los 50, el primer sintetizador de text to speech fue completado en el año 1968, y en el 61 el físico John Larry uso

un IBM 704 para sintetizar el habla, ese fue un evento bastante prominente en la historia de los laboratorios Bell de acuerdo con (Lambert, 1992).

Esta tecnología toma el texto escrito y lo pasa a sonido, para esto una voz ya pregrabada que está almacenada en una base de datos, se utiliza para entrenar un computador y crear una voz sintética, que se utilizará para leer texto escrito (Skinner, 2020).

Estado Del Arte

A Nivel Internacional

Título: The Impact of the Virtual Assistant (VA) on Language Classes (El impacto del asistente virtual en las clases de idiomas)

Autor: Ana Mihaela

Año: 2019

Lugar: Romanian-American University, Bucharest, Romania

Resumen: Este artículo trata sobre estudios acerca del uso de la inteligencia artificial en el salón de idiomas, donde las máquinas pueden ser equipadas con una inteligencia similar a la humana. Específicamente hablando las máquinas pueden ser entrenadas para el autoaprendizaje. De este modo, asistentes virtuales como Siri podrían convertirse en profesores sustitutos en las clases de idiomas. Una de las actividades que más consume tiempo es la interacción individual con los estudiantes, pero con una inteligencia artificial, se podría lograr que el estudiante haga pareja con un VA, y de alguna forma esta clase sea personalizada creando una conexión personal con el estudiante. También se convertiría en un entrenador, ayudando a aprender gramática y vocabulario, así como practicar conversaciones (Mihaela, 2019).

A Nivel Nacional

Título: Arquitectura de un asistente virtual para la producción de textos en la enseñanza y aprendizaje de idiomas

Autores: Diana C. Rivera, Liliana E. Machuca-Villegas

Año: 2014

Lugar: Universidad del Valle, Cali, Colombia

Resumen: Este artículo presenta la arquitectura de un Asistente virtual que ofrece funciones de apoyo al usuario en el proceso de producción de textos para la enseñanza y el aprendizaje de idiomas. Se definen dos funciones específicas para el comportamiento del asistente: la asistencia de escritura y la asistencia gramatical. En la asistencia de escritura se le presenta al estudiante recomendaciones de acuerdo al tipo de párrafo que desea escribir, y en la asistencia gramatical se le brinda al estudiante la corrección ortográfica y sintáctica de oraciones de los párrafos o textos que produce. El asistente integra características de agentes de software, ambientes virtuales de aprendizaje y elementos de la lingüística computacional que permiten realizar el procesamiento del lenguaje y dar recomendaciones a los estudiantes sobre su proceso de escritura. El principal propósito de este asistente es ayudar a los estudiantes a mejorar sus habilidades de escritura cuando interactúen en un entorno de aprendizaje virtual. La arquitectura propuesta se ha implementado en la plataforma denominada Lingweb, prototipo de un Sistema de Gestión de Cursos para la enseñanza y aprendizaje de idiomas (Rivera & Machuca-Villegas, 2014).

Marco Teórico

Aprendizaje Del Lenguaje

El lenguaje es una forma estructurada que usan los seres humanos para transmitir un mensaje, un sentimiento, una información etcétera, por tal motivo es necesario que el humano aprenda un lenguaje. Este proceso no sigue totalmente los pasos que se utilizan en el aprendizaje normal. Cuando se habla de aprendizaje, se busca entender y memorizar un saber, y el hombre tiene muchos de estos saberes debido al aprendizaje, pero no todo lo que él se sabe lo ha aprendido, sino que hubo procesos que lo llevaron a adquirir conocimientos, por ejemplo: el lenguaje nativo de los seres humanos fue adquirido, y después se aprendió a hablarlo y a escribirlo. La adquisición es un proceso que ocurre de forma subconsciente, es decir el humano no se da cuenta de ello, es algo que ocurre de forma inconsciente, por el cual la persona absorbe el sentido y significado de las palabras de un idioma, sin que él esté al tanto de que eso está pasando dentro su cerebro. Después de que la persona adquiere el idioma entonces lo aprende a hablar, ósea aprende a realizar los fonemas que dicho idioma tiene.

En este campo hay algunos misterios, de cómo esto sucede, pues cuando el humano nace, no tiene un libro de gramática o un traductor, pero lo que si todos tienen son oídos y ojos, y a través de estos puede escuchar lo que otros dicen e ir relacionando poco a poco los sentidos y sonidos con objetos, emociones y acciones. Es entonces donde la comunicación juega un rol importante, porque si no hay comunicación, no va a haber una buena adquisición del lenguaje, pues el desarrollo del lenguaje en los niños es impactado, cuando ellos escuchan una conversación y hacen parte de ella (How Do We Learn To Talk?, 2019).

Asistentes Virtuales

Los asistentes virtuales son aplicaciones que intentan imitar algunos de los comportamientos humanos, con el fin de interactuar y realizar tareas ordenadas por ellos. Estos sistemas pueden ayudar tanto a empresas que presten algún tipo de servicios como al usuario de los servicios. Estas herramientas están presentes en una gran variedad de dispositivos electrónicos como celulares y computadores de acuerdo con (Louise, 2020a), también en dispositivos que hacen parte del internet de las cosas.

Su funcionamiento se da por unos algoritmos de Entendimiento del Lenguaje Natural (NLU) que usan análisis semántico (Kavlakoglu, 2020) que le permiten determinar la intención del usuario. Cuando se crea un asistente virtual, normalmente se programa para realizar cierto tipo de actividades específicas, como ejemplo la consulta de la cita médica de la siguiente semana, también para almacenar recordatorios; así el programa está orientado con el objetivo de cubrir ese tipo de necesidad. Cuando el usuario da una consulta el asistente comienza a analizar el texto con un componente de la AI que se conoce como Procesamiento de lenguaje Natural (NLP) (Guerrini, 2016), el cual procesa el texto y busca las intenciones y sentimientos del usuario (Education, 2021). La consulta pasa por tareas de NLP como reconocimiento del habla que es el text to speech, también el etiquetado gramatical que es la identificación del frases o palabras según el uso que tengan o el contexto, otra de estas tareas es la desambiguación del sentido de la palabra, que permite determinar el sentido de alguna palabra que cuente con múltiples significados. Añadido a esto se tiene un proceso que permite encontrar entidades y se llama reconocimiento de entidades nombradas, Adicionalmente a estos procesos hay otro que permite la identificación de cuando dos palabras hacen referencia a una misma entidad y asimismo está el análisis sentimental que es el que intenta extraer

cualidades, actitudes emociones, sarcasmo, confusión y sospecha del texto, todo esto antes que el asistente entienda lo dicho por el usuario. Cuando el asistente ha entendido de qué está hablando el usuario, se pasa a una serie de condiciones preprogramadas en la sección de los diálogos, los cuales darán paso a las respuestas que el asistente brindará al usuario (IBM Cloud Education, 2020a).

Tecnología Text To Speech

Esta herramienta es la que permite que los sistemas puedan generar audio de textos, esto sucede por unas operaciones que realizan estos sistemas. La normalización es uno de estas operaciones, la cual se encarga de reducir la ambigüedad que se pueda presentar en la lectura del texto. Hay ciertos aspectos en un texto que pueden causar ambigüedades como las fechas, acrónimos, horas etcétera; en este punto se busca determinar la correcta pronunciación para determinada palabra, numero o carácter, y para esto se usan los Modelos Escondidos de Markov, los cuales ayudan a determinar la posible pronunciación de dichos elementos con probabilidades, teniendo en cuenta los contextos de la frase. En el momento en que sea reducido la ambigüedad otros procesos entran para estructurar la sentencia que va a ser leída, el sistema va a comenzar a tomar lo fonemas y los va a alinear con su carácter respectivo. Esta parte es una de las más complicadas, ya que algunas palabras no se leen como se escribe, y algunos idiomas tiene un cambio completo en la pronunciación respecto de los caracteres usados en determinada palabra. Para estos casos se establecen reglas que el sistema usará para alcanzar una pronunciación más exacta del texto. Después de que el sistema determino cuales son los fonemas a utilizar, tomará los sonidos grabados por humanos o generados por computadora para leer en voz alta lo escrito.

Los sonidos grabados por humanos son lo más comunes en los asistentes virtuales, su sonido es natural, donde se toman frases grabadas de humanos y se seccionan al punto de tomar fonemas de las palabras que fueron pronunciadas. Con estos fonemas individuales el sistema puede construir palabras que nunca antes le han sido dadas concatenando los segmentos.

Los sonidos generados por computadora o sintetizadores del habla, tienen una cobertura más amplia que los concatenados, al poder pronunciar cualquier tipo de palabras y fonemas, no tienen limitación de idioma, y con cambios en la frecuencia puede sonar masculino o femenino, pero su desventaja es que tienden a sonar como un robot (Woodford, 2021).

Tecnología Speech To Text

La tecnología de Speech To Text, comprende una gran variedad de procesos, los cuales comienzan cuando la voz es grabada, el audio es enviado a un servidor o computador el cual con el componente de modelo acústico fraccionará el audio en pequeñas unidades, y lo comparará con una base de datos, con cientos de fonemas, estos modelos usan algoritmos entrenados de deep-learning (Voximplant, 2020). Cuando las unidades de audio fraccionadas comienzan a tener una probabilidad alta con respecto del fonema es entonces cuando se forman las palabras. De allí otra serie de operaciones entran a aplicar reglas gramaticales, los Modelos Escondidos de Markov (Amberscript, 2019) con su componente lingüístico, que buscan determinar cuál es la posible palabra dentro de la oración, en este escenario los contextos son cruciales, especialmente cuando las palabras suenan iguales pero su sentido o escritura es distinto, porque ayudan a determinar con gramática, cual es la palabra que corresponde con el sonido de la voz grabada.

Hay diferentes tipos de reconocimiento del habla, están los basados en el modelo del hablante como los modelos dependientes del hablante y los independientes del hablante. También están los basados en declaraciones, que se basan en palabras aisladas, palabras conectadas, habla continua y habla espontánea. Otro tipo de reconocimiento es el basado en el vocabulario (Saksamudre et al., 2015).

Dentro de los modelos o métodos más usados para el reconocimiento del habla están los ya mencionados Modelos Escondidos de Markov, también el Procesamiento de Lenguaje Natural, las redes neuronales, N-gram y Speaker Diarization (IBM Cloud Education, 2020b).

Metodología De Desarrollo

Esta metodología híbrida se desarrolló para la elaboración de este proyecto, la cual toma algunas de las fases de metodologías Scrum y Kanban; donde se implementan de una consecución de tareas de inicio a fin, también se tienen planificaciones donde se determina lo que hay que hacer, lo que se está haciendo y lo que está hecho.

Scrum:

Scrum es una metodología de desarrollo del software muy utilizada que se basa en el trabajo en equipo con consecuentes Sprints. De modo que, Scrum es una metodología ágil, que tiene como objetivo controlar, organizar y planificar proyectos con el fin de hacer más eficiente la creación y entrega de software. En esta metodología se puede planificar los Sprints que van de una semana a cuatro semanas y cuando se termina el Sprint, se revisa el trabajo validado de la

anterior semana. En función de esto, se priorizan y planifican las actividades en las que se invierten los recursos en el siguiente Sprint. La metodología Scrum se centra en ajustar sus resultados y responder a las exigencias reales y exactas del cliente. De ahí, que se vaya revisando cada entregable, ya que los requerimientos van variando a corto plazo. Entre las principales características de la metodología Scrum, destaca que es un desarrollo incremental en lugar de la clásica planificación del desarrollo completo de un producto o servicio. Los equipos Scrum se caracterizan por ser autoorganizados. Y se centra en el producto final y en la calidad del mismo” (Sinnaps, 2020) además de esto la metodología Scrum en su estructura tiene reuniones que se hacen todos los días donde como se mencionó anteriormente se prioriza lo que hace falta por entregar.

Kanban:

Kanban es una metodología que se ha venido utilizando desde hace mucho tiempo, más o menos desde hace más de 50 años. Está enfocada en la eficiencia, donde se planean las actividades y sobre una tabla se ubican tres columnas, donde en la primera se pondrá lo que está para realizarse, en la segunda se ubican las actividades que se están realizando, y la tercera van las que ya están hechas. Cuando las tareas están siendo realizadas aparecen tres estados, uno es trabajando, otro en esperando y por último en revisión (kanbanize, 2017).

Metodología Híbrida:

La metodología híbrida se ha desarrollado para este proyecto y está compuesta de tres fases y siete sub fases. En la fase inicial está dividida en varias sub fases las cuales son: análisis e investigación, diseño y planificación. El análisis y la investigación de esta fase es la que se

encarga de tomar los requerimientos, de investigar sobre el producto final deseado por el cliente, y de analizar todas las opciones posibles para desarrollar el prototipo. La siguiente etapa, que partir de la toma de requerimientos se realiza un diseño o un prototipo en borrador, el cual será el punto o puerto hacia donde ira el desarrollo del proyecto. En la planificación se toman los tiempos que se van a utilizar para el desarrollo de este proyecto, en estos tiempos va lo que es el proceso y la implementación, la fecha de entrega, y dónde también se definen que lo primero son los requerimientos.

En la segunda fase de la metodología se tiene el proceso, que es la sección donde se va aplicar lo que se planificó en la fase anterior, se presentan como sub fases el desarrollo y el testeo. Aquí es donde se desarrolla la codificación y el diseño que fue puesto en la lista de tareas. Estando en esta etapa si en el desarrollo necesita una rectificación de alguna cosa entonces se aplica el cambio necesario haciendo un rediseño del prototipo. Después de terminar el prototipo se pasa a la sub fase del testeo, la cual es repetitiva y retroalimenta al diseño. Si se cumple con las expectativas pasa la siguiente fase, sino vuelve a la parte del desarrollo donde se retroalimenta, y se corrigen todas aquellas cosas que no pasaron en la fase del testeo. Cuando finalmente termina el testeo con un nivel que sea de aceptación por parte de los integrantes del proyecto, tanto clientes como desarrolladores, pasa a la tercera fase que es la fase final donde se pone en implementación y en operación el prototipo desarrollado, si surgen nuevos requerimientos, y hay un acuerdo de nuevo prototipo se reinicia todo el proceso dando como resultado un nuevo prototipo en el tiempo.

Fase 1. Inicial

- Análisis e investigación,

- Diseño (prototipo final en borrador),
- Planificación (tiempo, división de cargas (lista de tareas), primero los requerimientos).

Fase 2. proceso

- Desarrollo (codificación y diseño (lista de tareas), reuniones de avance (se revisa lo que he hecho (lista de tareas), lo que estoy haciendo y lo que falta), replanificación si es necesaria),
- Testeo. (Fase repetitiva, se retroalimenta diseño, pero si cumple con las expectativas pasa a la siguiente fase).

Fase 3. Final

- Implementación,
- Operación. (Si surgen nuevos requerimientos se reinicia en todo el proceso).

Implementación De La Metodología De Desarrollo

La metodología se aplicó con la investigación de todos los elementos necesarios para planificar el desarrollo del producto final, también determinó lo que se iba a realizar, y cuáles eran los tiempos que se iban a manejar. Al principio la búsqueda estuvo orientada al conocimiento de los asistentes virtuales, añadido a esto se adquirieron softwares para el desarrollo del asistente virtual, como el software de Unity, el gestor de bases de datos SQLite y la creación de la cuenta en IBM Cloud para creación del servicio del asistente y los servicios de Text to Speech y Speech to Text. Después se realizó un prototipo en borrador, para establecer un marco de referencia y a partir de allí se realizó la planificación de acuerdo a las actividades a

realizar. Siendo así se inició el desarrollo del proyecto con la creación de intenciones en el asistente IBM Watson, cumpliendo los dos primeros objetivos específicos.

El objetivo número tres, es el desarrollo de la aplicación en Unity, donde se crearon escenas para el inicio de la aplicación y cada una de las secciones; también se instalaron los respectivos componentes para el funcionamiento del software como el SDK de Watson y el plugin de SQLite.

El código de la aplicación en parte es propio, pero específicamente las secciones de Text to Speech, el Assistant y Speech to Text tuvieron como base los ejemplos que da IBM Watson en su SDK, también se tomaron ejemplos del proyecto programación de citas Oculus Lipsync Main anteriormente hecho por Scott Hawang, y se adaptaron a este proyecto.

Capítulo 2: Análisis de tecnologías para desarrollo de asistentes inteligentes

Tecnologías para diseño de aplicaciones con asistentes virtuales

Las tecnologías para el desarrollo de asistentes virtuales inteligentes abundan, pero hay unas herramientas tecnológicas que están más avanzadas que otras. Dentro de los lenguajes se encuentran C++, Java y Python el cual cuenta con varias librerías para desarrollo de herramientas de inteligencia artificial.



Ilustración 1 Languages de programación usuados en la inteligencia artificial

Fuente: Los logos fueron tomados de (Oracle, n.d.) y (Python, n.d.) y editadas en Gimp

El logo de C++ es propio hecho en Gimp

Unity

Para este proyecto se utilizará Unity qué es un entorno de desarrollo enfocado en video juegos, pero también se puede crear otro tipo de aplicaciones. Así mismo exporta a diferentes tipos de plataformas como web, Android, IOS y Windows. Unity es una herramienta de uso libre, claro está que sí la ganancia es igual o mayor a 100 mil dólares anuales habrá que pagar una cantidad para seguir utilizando la plataforma, y cuyo uso puede llegar a ser fácil con algún tiempo de práctica(Sinicki, 2021).

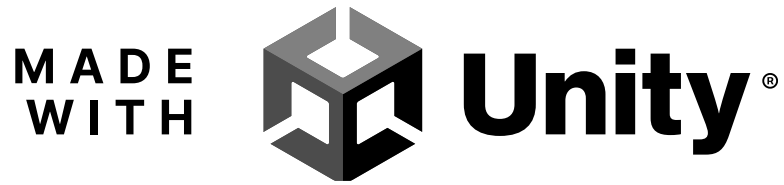


Ilustración 2 Unity herramienta de desarrollo de video juegos

Fuente: (Unity, n.d.)

C#

El lenguaje de programación usado será Csharp, que es uno de los lenguajes utilizados por defecto en Unity. Este lenguaje pertenece a Microsoft y permite crear aplicaciones de escritorio entre otros softwares. Gracias a las librerías que integra Unity muchas de las funciones ya están listas y lo único que se debe hacer es invocarlas.



Ilustración 3 Csharp

Fuente: (Microsoft, n.d.)

SQLite

Los datos serán almacenados de SQLite, un gestor sencillo, el cual puede ser fácilmente integra en el proyecto con Unity y C#.



Ilustración 4 Gestor de base de datos SQLite

Fuente: by author D. Richard Hipp to the public domain.

Watson Assistant, Speech to Text y Text to Speech

Se utilizarán los servicios de Watson Assistant, Speech to Text y Text to Speech, los cuales se puede integrar con las herramientas mencionadas anterior mente. Estos servicios pueden ser utilizados de forma gratuita, siempre y cuando no se exceda el plan lite que ofrece IBM Cloud. Cabe mencionar que por su versatilidad el asistente de IBM, puede llegar a ser sencillo de programar, pero a mayor complejidad en el dialogo, mayor complejidad en la programación del mismo.

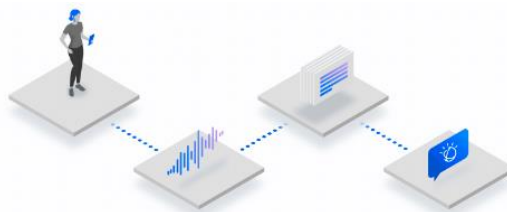


Ilustración 5 Servicios de Watson Assistant, Speech to Text y Text to Speech

Fuente: (IBM, n.d. -a)

Capítulo 3: Definición de intenciones enfocadas en el desarrollo conversacional

Intenciones

Para este asistente es necesario definir intenciones que él pueda reconocer dentro de las expresiones que el usuario utiliza, estas intenciones están enfocadas en dos tópicos principales que el asistente va a manejar, las cuales son fútbol y carros. También se crearán intenciones orientadas a frases que incluyan la acción de dar gracias, así mismo de saludar, despedirse, manifestar gustos o disgustos. Igualmente, hay intenciones que están vinculadas a declaraciones usadas por el usuario cuando manifiesta que ignora algún conocimiento o que no lo recuerda, de tal manera que el asistente pueda entender lo que el usuario dice. Estas intenciones son:

- **Bad:** Expresiones que describen algo feo o desagradable.
 - Ej: It was a horrible experience
- **Bad_player:** Expresiones que describen cuando el usuario no es un buen jugador.
 - Ej: I am a bad player
- **Best_player:** Expresiones que describen cuando se habla de algún buen jugador.
 - Ej: I think that Messi has been the best player in history
- **Confirmation:** Expresiones que describen una afirmación.
 - Ej: Absolutely, of course, ok, yeah
- **Conversation:** Expresiones que denotan la intención de querer hablar.
 - Ej: Could I speak with you
- **Do_not_know:** Expresiones que denotan la falta de conocimiento frente alguna pregunta.
 - Ej: I don't know

- Do_not_like: Expresiones que denotan el no gusto por alguna cosa.
 - Ej: I don't like it
- Do_not_play_often: Expresiones donde el usuario dice que no juega.
 - Ej: I don't play
- Do_not_remember: Expresiones que el usuario dice cuando no se acuerda.
 - Ej: I don't remember
- Do_not_understand: Expresiones que el usuario dice cuando no entiende.
 - Ej: I don't understand
- Dont_like_game: Expresiones que el usuario dice cuando no le gusta algo que pase dentro de un partido.
 - Ej: In a game where people start fighting each other
- Favorite: Expresiones que el usuario dice cuando tiene cierta preferencia.
 - Ej: I love cars, I love the blue color
- Feel_normal: Expresiones que el usuario dice cuando no es entusiasta o desilusionado sobre algo.
 - Ej: Nothing extraordinary
- FeelBad: Expresiones que el usuario dice cuando no le gusta algo o le hace sentir mal.
 - Ej: Bad, don't like it
- FeelGood: Expresiones que el usuario dice cuando algo lo hace sentir bien.
 - Ej: It feels good
- General_About_You: Expresiones que el usuario dice cuando pregunta información del chatbot.
 - Ej: Who are you?

- General_Ending: Expresiones que el usuario dice cuando quiere finalizar la conversación.
 - Ej: Bye, see you later
- General_Human_or_Bot: Expresiones que el usuario dice cuando pregunta con quien está hablando.
 - Ej: Are you a human?
- General_Jokes: Expresiones que el usuario dice cuando quiere que le digan una broma.
 - Ej: tell me a joke.
- General_Positive_Feedback: Expresiones que el usuario dice cuando con cuerda con algo o piensa que está bien.
 - Ej: This is great, this is good
- GiveThanks: Expresiones que el usuario dice cuando está agradecido.
 - Ej: Thank you, thanks
- Good: Expresiones que el usuario dice cuando con cuerda con algo o piensa que está bien.
 - Ej: Amazing, fantastic
- Good_player: Expresiones que el usuario dice cuando cree que es un buen jugador.
 - Ej: I am good player
- Greeting: Expresiones que el usuario dice cuando quiere saludar.
 - Ej: Hi, hello
- How_often_play_football: Expresiones que el usuario dice cuando da una aproximación de los tiempos que juega football.
 - Ej: Almost everyday
- I_dont_have: Expresiones que el usuario dice cuando no tiene algún objeto o cosa.

- Ej: I don't really have
- I_have_played: Expresiones que el usuario dice cuando menciona juegos que ha realizado en el pasado.
 - Ej: I have played football
- I_like_it: Expresiones que el usuario dice cuando algo le gusta.
 - Ej: I like it
- I_think: Expresiones que el usuario dice cuando cree o piensa que algo es de una u otra manera.
 - Ej: I guess is Neuer
- JokeConversation: Expresiones que el usuario dice cuando pide unos momentos antes de hablar.
 - Ej: Please give a sec
- Knowing: Expresiones que el usuario dice cuando quiere saber sobre qué temas puede hablar con el chatbot.
 - Ej: What you know
- Knowledge: Expresiones que el usuario dice cuando habla o menciona la palabra conocimiento.
 - Ej: Knowledge, know
- Laugh: Expresiones que el usuario usa cuando se rie.
 - Ej: huhuhuhu
- Like_Drive: Expresiones que el usuario dice cuando manifiesta que le gusta manejar algún vehículo.
 - Ej: I like to drive

- Like_Football: Expresiones que el usuario dice cuando le gusta el football.
 - Ej: I like football
- Like_Run: Expresiones que el usuario dice cuando le gusta andar a altas velocidades.
 - Ej: I like to hit the gas
- Like_slow: Expresiones que el usuario dice cuando les gusta andar lento.
 - Ej: I like to go slowly
- Negation: Expresiones que el usuario dice cuando niega o dice que no.
 - Ej: No, negative
- Never_played: Expresiones que el usuario dice cuando nunca ha jugado.
 - Ej: I don't play
- Problem_on_road: Expresiones que el usuario dice cuando tuvo inconvenientes en carretera.
 - Ej: We ran out of fuel
- So_so: Expresiones que el usuario dice cuando habla en términos medios.
 - Ej: So so, more or less, a bit
- Whatever: Expresiones que el usuario dice cuando dice lo que sea.
 - Ej: Whatever
- Why_I_like_road_trip: Expresiones que el usuario dice cuando no se acuerda.
 - Ej: I like that smell of nature
- Why_like_loud_car: Expresiones que el usuario dice cuando explica porque es que le gustan los carros ruidosos.
 - Ej: I like the sound

Capítulo 4: Desarrollo de la aplicación

Asistente

Descripción Del Proceso De Comunicación Con Watson Assistant

El proceso de comunicación del Asistente de IBM Watson es realizado con inteligencia artificial (Virtual Agent - IBM Watson Assistant | IBM, 2018), donde se utilizan una serie de algoritmos estadísticos para determinar las intenciones del usuario. Las frases son analizadas en búsqueda de intenciones, las cuales establecerán las acciones que el asistente irá a realizar. Esto es efectuado por el Procesamiento del Lenguaje Natural (NLP) de acuerdo con (Introduction to Watson Assistant - IBM Developer, 2019) que es una rama de la inteligencia artificial. Esta tecnología analiza el habla o texto humano en búsqueda de sentido (Laura, 2021), de tal manera que el asistente pueda dar una respuesta acertada.

El Procesamiento del Lenguaje Natural se encuentra en distintos servicios como GPS operados por voz, Asistentes digitales, softwares de dictado (IBM Cloud Education, 2020) entre otras tecnologías, donde un subproceso conocido como Entendimiento del Lenguaje Natural (NLU) toma acción. Este proceso es el que busca las intenciones dentro del mensaje dado por el humano, y para manejar las respuestas el asistente usa (DM) Dialogue Manager, el cual con una serie de diálogos y condiciones previamente incluidas responde a lo que le ha sido preguntado (Louise, 2020b).

La aplicación en el primer y segundo nivel usa solo el servicio de Speech to Text, el cual es iniciado cuando se inician cada uno de estos niveles. En el tercer nivel inicia los servicios, de

Speech to Text, Text to Speech e IBM Watson, los cuales comienzan a correr, después de esto el asistente saluda, y la aplicación manda el saludo al servicio de Text to Speech, el cual lo convierte a audio y lo envía de vuelta a la aplicación para ser reproducido en los parlantes. La aplicación evalúa si hay algo más para reproducir, sino da paso al usuario abriendo el micrófono. Cuando el usuario comienza a hablar, la aplicación comienza a grabar y a pasar el mensaje al servicio Speech to Text, que retornará un texto digital. Hecho esto la aplicación remitirá el texto digital al asistente hecho IBM Watson, quien buscará las intenciones y responderá a la aplicación con un texto digital. La aplicación enviará el texto digital al servicio Text to Speech el cual pasará el texto digital a audio digital, que irá a la aplicación para ser reproducido, como se muestra en la ilustración 6.

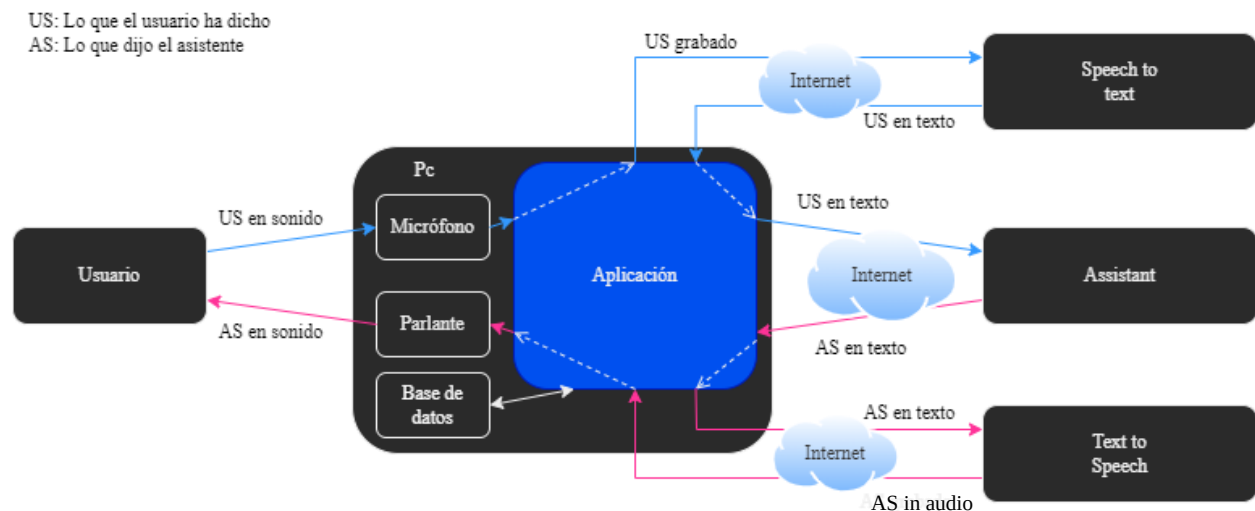


Ilustración 6 Flujo de conversación con el asistente

Estructura de la conexión de la aplicación con los distintos componentes utilizados en la misma, la aplicación se conecta a la base de datos, también al micrófono y al parlante, para poder acceder a los servicios de IBM Cloud se conecta a los servidores.



Ilustración 7 Conexión de la aplicación con los componentes que utiliza

Diagramas UML

Representación del uso de la app a través de diagramas UML

Diagrama General de Actividades para la Práctica de Conversación

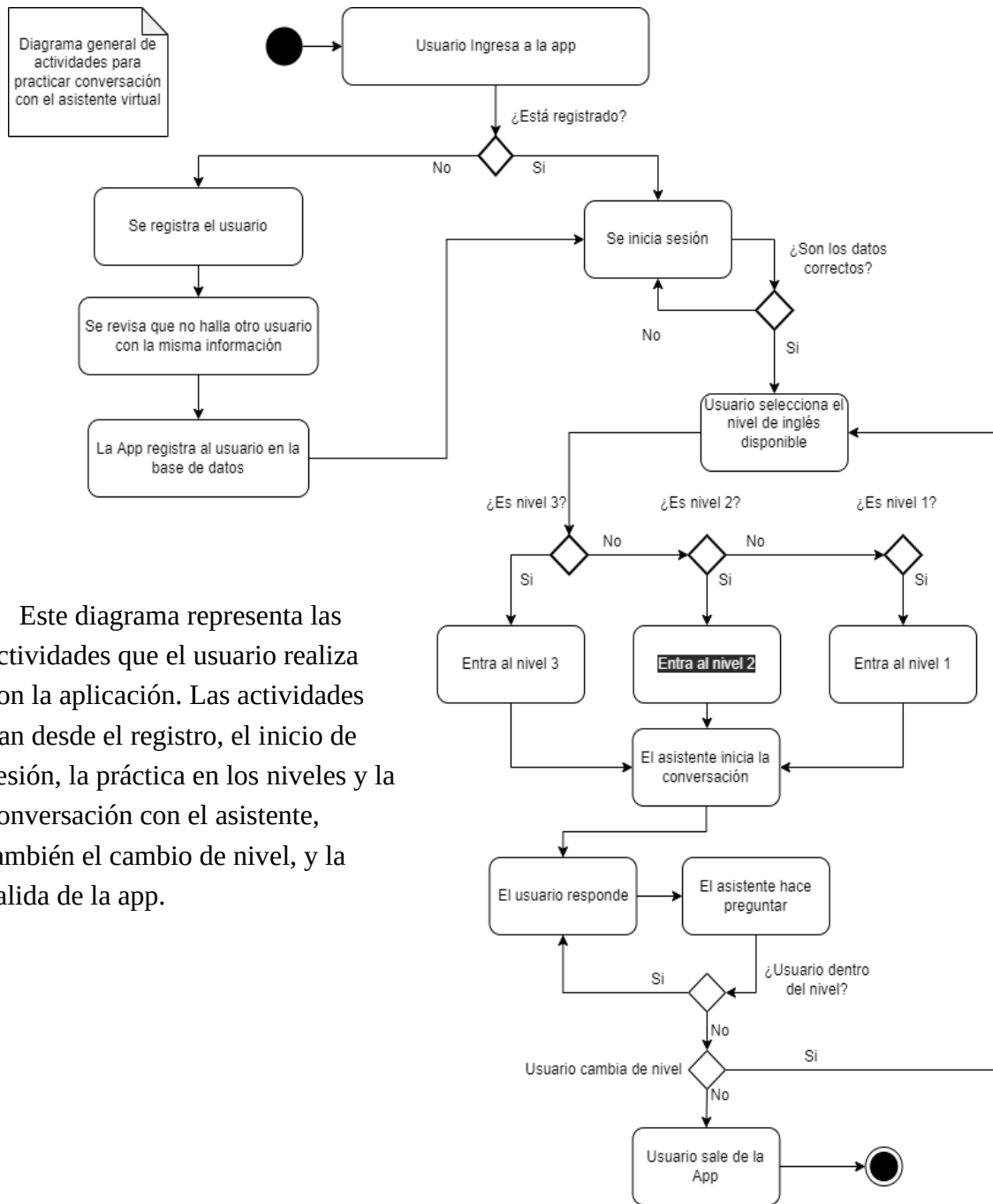


Diagrama 1: Diagrama de actividades para la práctica de Conversación

Diagrama de Estados

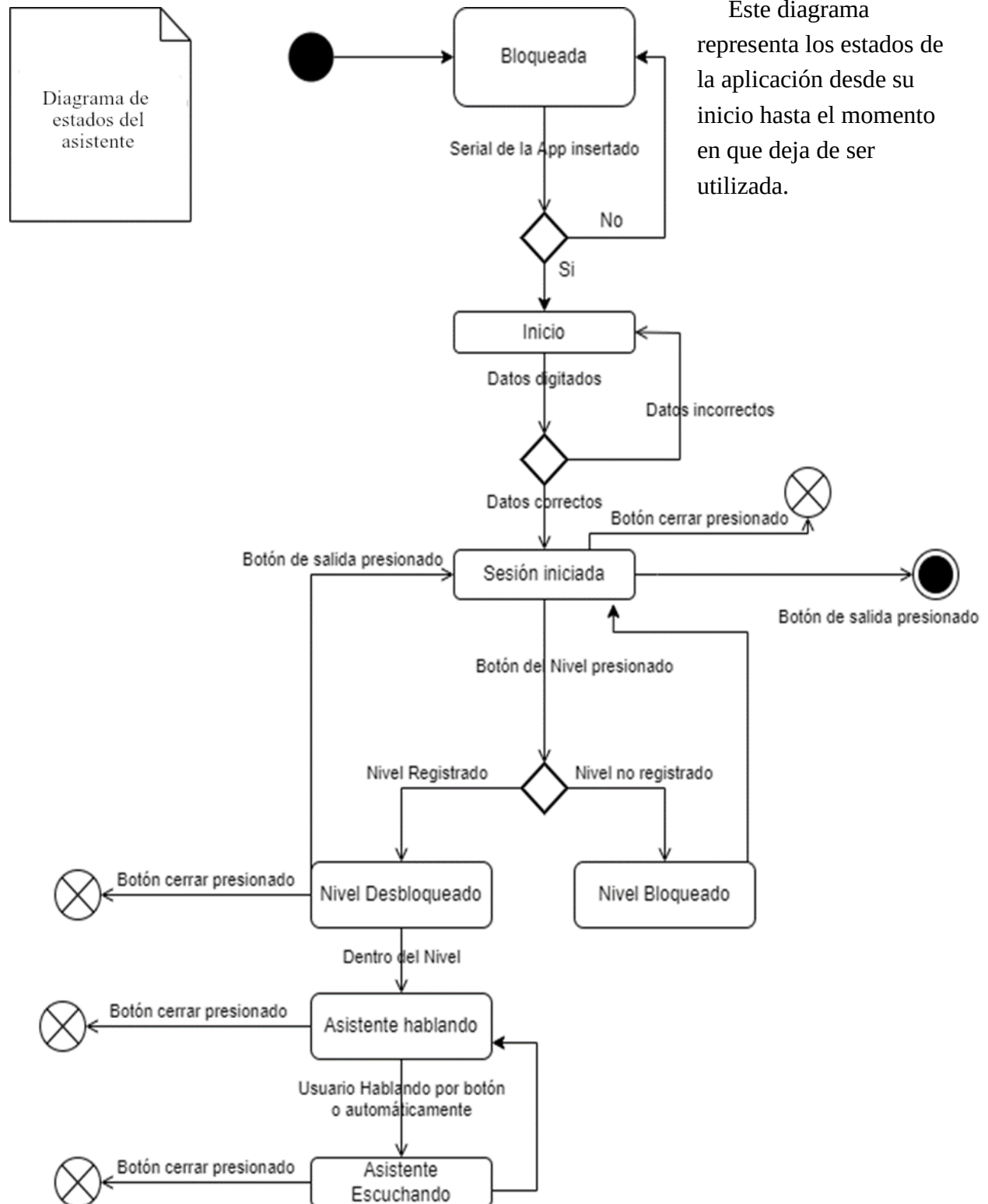


Diagrama 2 Diagrama de estado de un usuario entrado a practicar con el asistente

El diagrama de casos de uso de la aplicación representa todas las interacciones que los actores (Usuario, Base de datos, Speech to Text, Text to Speech, Wastson Assistant) pueden tener con el sistema.

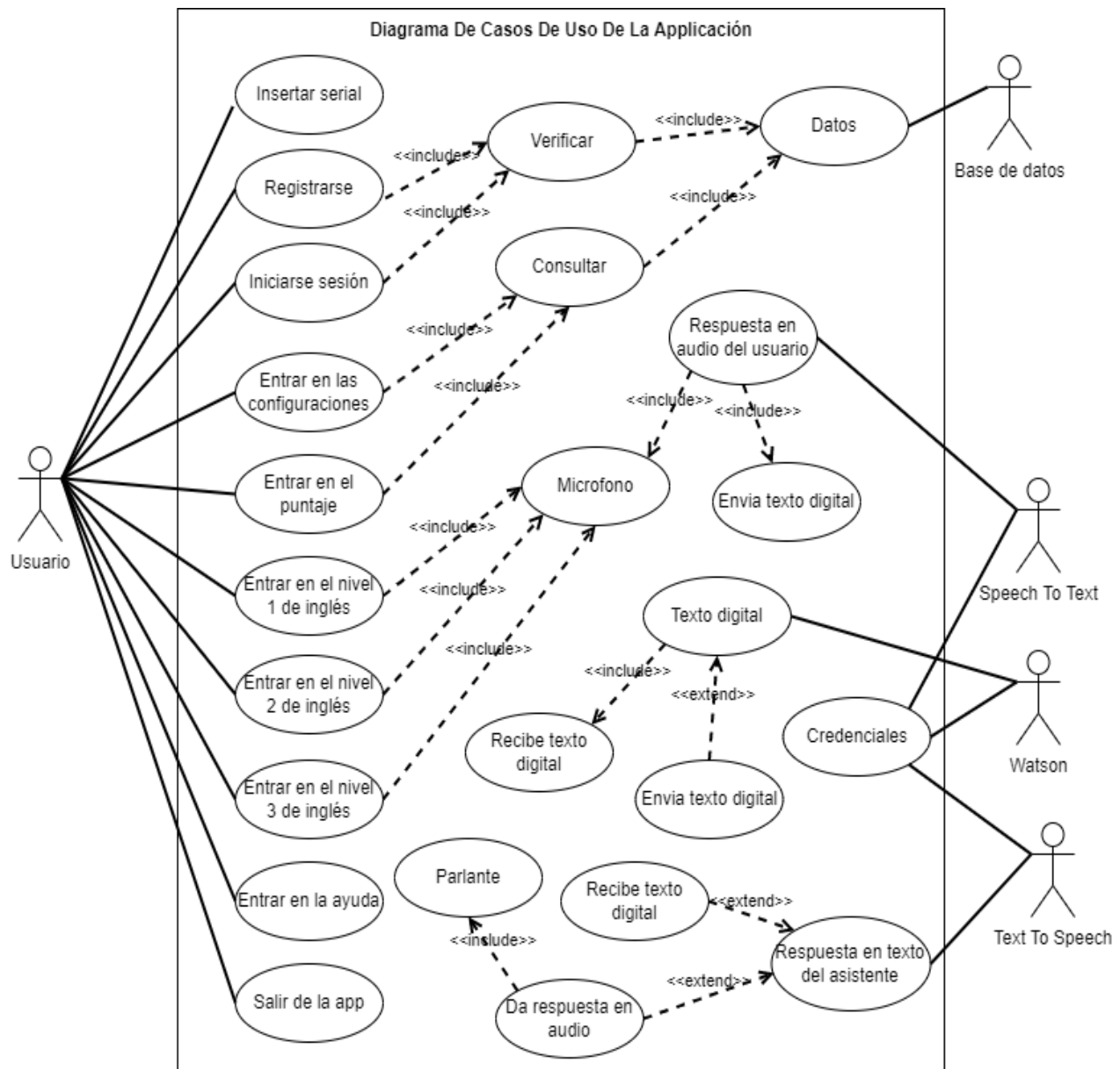


Diagrama 3 Diagrama general de casos de uso de la Aplicación

Diagrama de Casos de Uso Usuario

Este diagrama representa las interacciones específicas del usuario con el sistema

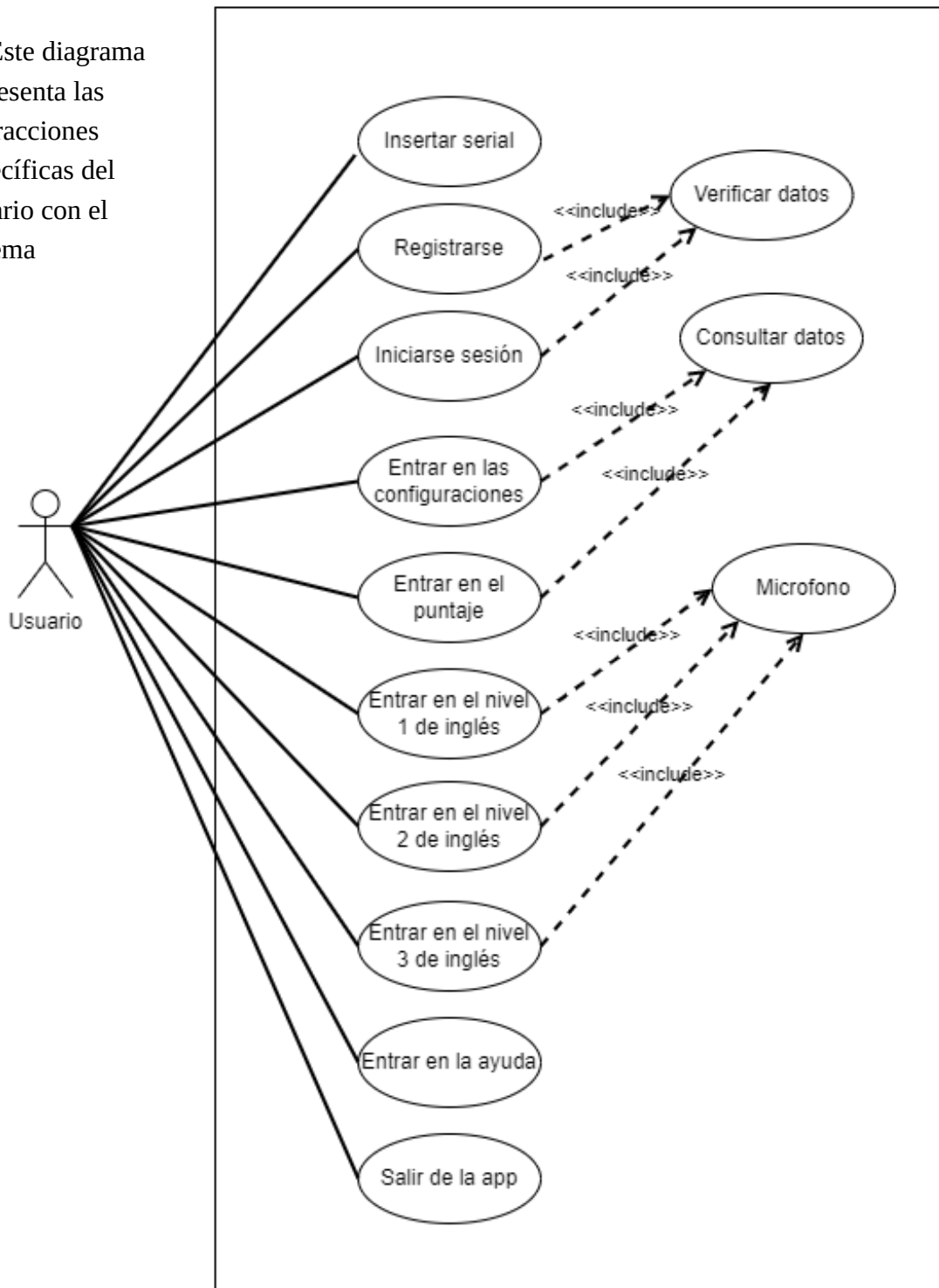


Diagrama 4 Diagrama de casos de uso Usuario

Diagrama de Casos de Uso Watson

Este diagrama representa las interacciones específicas del Watson Assistant con el sistema

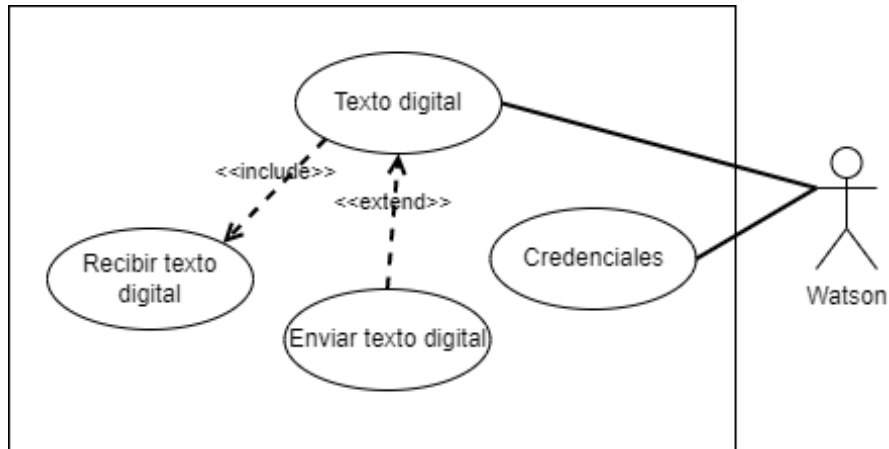


Diagrama 5 Diagrama de casos de uso Watson

Diagrama de Casos de Uso Speech to Text

Este diagrama representa las interacciones específicas del servicio Speech to Text con el sistema

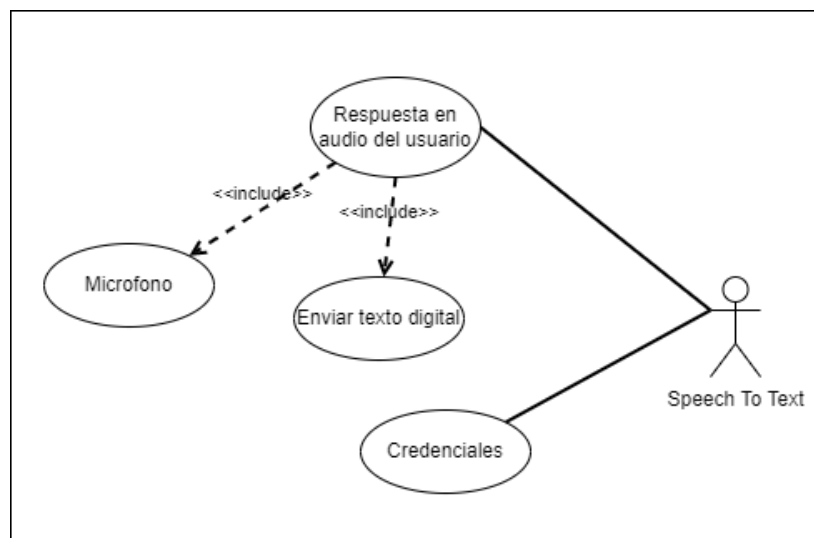


Diagrama 6 Diagrama de casos de uso Speech To Text

Diagrama de Casos de Uso Text to Speech

Este diagrama representa las interacciones específicas del servicio Text to Speech con el sistema

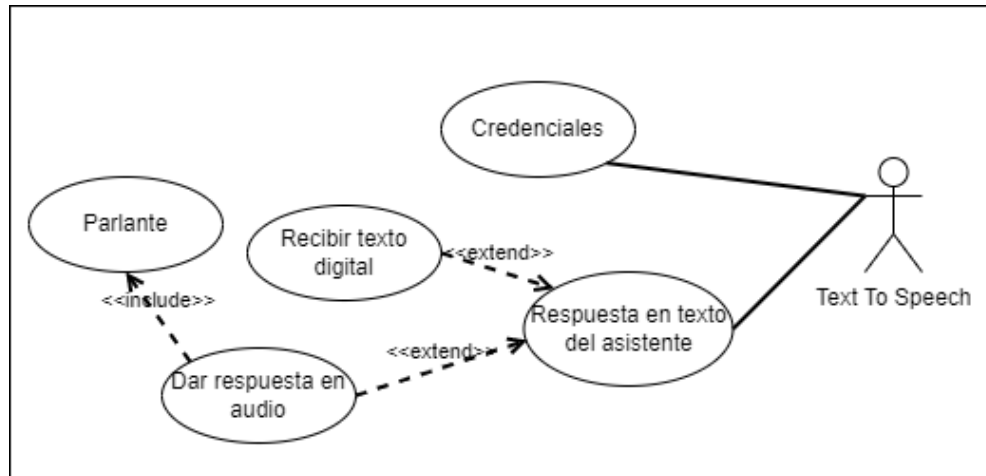


Diagrama 7 Diagrama de casos de uso Text to Speech

Diagrama de Casos de Uso Base de Datos

Este diagrama representa las interacciones específicas de la base de datos con el sistema

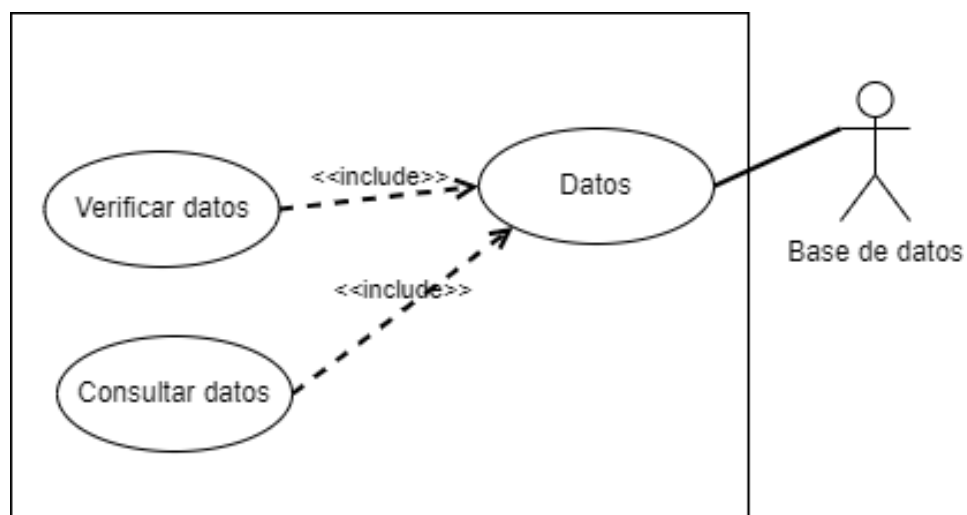


Diagrama 8 Diagrama de casos de uso Base de Datos

Identificación de actores

- Los actores del modelo son el usuario, la base de datos, IBM Watson Assistant, el servicio Text to Speech, el servicio Speech to Text y la Aplicación.

Identificación de actividades de cada actor

Los actores tienen actividades que realiza con la aplicación, la realización de estas actividades son las que permiten que la app cumpla con su propósito.

La Tabla 1 Actividades del usuario describe el rol que tiene el usuario con la aplicación, función que como usuario tiene y que el solo puede ejecutar.

Tabla 1 Actividades del usuario

AC-01	Usuario
Caso de uso	Insertar serial, iniciar sesión, entrar en el nivel de inglés, entrar en la ayuda, Entrar en las configuraciones, Entrar en el puntaje, Salir de la app, verificar datos, consultar datos, micrófono.
Descripción	El usuario es el encargado de registrarse y loguearse en la aplicación, así como de navegar por cada una de las opciones.

La Tabla 2 Actividades de IBM Watson Assistant describe las funciones que el asistente tiene como actor en el sistema.

Tabla 2 Actividades de IBM Watson Assistant

AC-02	IBM Watson Assistant
Caso de uso	Texto digital, recibir texto digital, enviar texto digital, credenciales.
Descripción	Watson Assistant analiza y responde de acuerdo con las intenciones que encontró en el texto digital

La Tabla 3 Actividades de Text to Speech describe las acciones que este servicio realiza en la aplicación

Tabla 3 Actividades de Text to Speech

AC-03	Text to Speech
Caso de uso	Respuesta en texto del asistente, recibir texto digital, dar respuesta en audio, parlante, credenciales.
Descripción	Pasa el texto digital a audio

La Tabla 4 Actividades de Speech to Text describe las acciones que este servicio realiza en la aplicación

Tabla 4 Actividades de Speech to Text

AC-04	Speech to Text
Caso de uso	Respuesta en audio del usuario, micrófono, enviar texto digital, credenciales.
Descripción	Pasa el audio a texto digital

La Tabla 5 Actividades de Base de datos describe las acciones que realiza la base de datos en la aplicación

Tabla 5 Actividades de Base de datos

AC-05	Base de datos
Caso de uso	Registros, verificar datos, consultar datos
Descripción	Contiene los usuarios que se han registrado en el dispositivo de forma local, también los puntajes y el nivel donde actualmente se encuentran.

Tabla De Descripción De Cada Caso De Uso

La Tabla 6 Descripción del caso de uso del usuario describe las posibles interacciones del usuario con el sistema, así como los objetivos, requisitos, secuencias y frecuencia.

Tabla 6 Descripción del caso de uso del usuario

AC-01	Usuario	
Autores	Yonathan Javier Jaimes Mantilla	
Objetivos Asociados	1. Practicar inglés en los niveles.	
Requisitos Asociados	1. El usuario debe registrarse para poder ingresar a la aplicación.	
Descripción	El usuario se registrará, y después de esto iniciará sesión, donde de acuerdo al nivel que haya escogido tendrá acceso a uno o los tres niveles de inglés, donde podrá practicar la conversación con el asistente.	
Precondición	Para ingresar a la aplicación debe registrarse en la misma, y solo podrá acceder a los niveles que sean menores o iguales al que haya elegido.	
Secuencia normal	Pasos	Acción
	1	Registrar
	2	Iniciar sesión
	3	Entrar en los niveles
	4	Practicar pronunciación
	5	Consultar puntaje
	6	Practicar con el asistente
	7	Actualizar datos
Postcondición	Ninguna	
Frecuencia	Varias ocasiones en la semana	
Comentarios	En caso de necesitar ayuda, intente revisar la sección de ayuda que la aplicación ofrece.	

La Tabla 7 Descripción del caso de uso de Asistente Watson describe las posibles interacciones del asistente de IBM con la aplicación, así como los objetivos, requisitos, secuencias y frecuencia.

Tabla 7 Descripción del caso de uso de Asistente Watson

AC-02	IBM Watson Assistant	
Autores	Yonathan Javier Jaimes Mantilla	
Objetivos Asociados	1. Reconocer las intenciones y entidades 2. Responder de manera apropiada al usuario.	
Requisitos Asociados	1. El asistente debe contar con las credenciales apropiadas.	
Descripción	Iniciará saludando al usuario, y esperando un saludo de vuelta, después hará un juego de preguntas, de esta forma se realizará la conversación con el usuario.	
Precondición	Ninguna	
Secuencia normal	Pasos	Acción
	1	Validar credenciales
	2	Iniciar conversación
	3	Saludar al usuario
	4	Buscar intenciones
	5	Buscar entidades
	6	Dar respuesta
	7	Realizar el juego la preguntas y respuesta con el usuario.
Postcondición	Ninguna	
Frecuencia	Varias ocasiones en la semana	
Comentarios	Ninguno	

La Tabla 8 Descripción del caso de uso del servicio Text to Speech describe las eventuales interacciones del servicio de Texto to Speech con el sistema, por consiguiente, describe objetivos, requisitos, secuencias y frecuencia entre otros.

Tabla 8 Descripción del caso de uso del servicio Text to Speech

AC-03	Text to Speech	
Autores	Yonathan Javier Jaimes Mantilla	
Objetivos Asociados	1. Convertir de texto a audio las respuestas enviadas por Watson Assistant.	
Requisitos Asociados	1. Este servicio debe contar con las credenciales apropiadas para poder funcionar.	
Descripción	Cuando el asistente haya encontrado una intención que fuere acorde con lo que el usuario quiso decir, enviar una respuesta texto digital, el cual será transferida a servicio de Text to Speech el cual pasará a audio el contenido del texto digital.	
Precondición	Ninguna	
Secuencia normal	Pasos	Acción
	1	Validar credenciales
	2	Convertir el text a audio
Postcondición	Ninguna	
Frecuencia	Varias ocasiones en la semana	
Comentarios	Ninguno	

La Tabla 9 Descripción del caso de uso del servicio Speech to Text describe las potenciales interacciones del servicio Speech to Text con el programa, así como objetivos, requisitos, secuencias y frecuencia.

Tabla 9 Descripción del caso de uso del servicio Speech to Text

AC-04	Speech to Text	
Autores	Yonathan Javier Jaimes Mantilla	
Objetivos Asociados	1. Convertir de audio a texto las respuestas que da el usuario.	
Requisitos Asociados	2. Este servicio debe contar con las credenciales apropiadas para poder funcionar.	
Descripción	Después que el asistente haya hablado se activará el micrófono para darle paso al usuario, el cual dará una respuesta a la pregunta, esta respuesta será convertida en un texto digital por el servicio de Speech to Text.	
Precondición	Ninguna	
Secuencia normal	Pasos	Acción
	1	Validar credenciales
	2	Convertir el audio a texto
Postcondición	Ninguna	
Frecuencia	Varias ocasiones en la semana	
Comentarios	Ninguno	

La Tabla 10 Descripción del caso de uso de la Base de Datos describe las eventuales interacciones del base de datos con el sistema, así como sus objetivos, requisitos, secuencias y frecuencia.

Tabla 10 Descripción del caso de uso de la Base de Datos

AC-05	Base de datos	
Autores	Yonathan Javier Jaimes Mantilla	
Objetivos Asociados	1. Almacenar la información pertinente al usuario.	
Requisitos Asociados	Ninguna	
Descripción	La base de datos almacenará los datos del usuario, así como los puntajes ganados en los niveles 1 y 2.	
Precondición	Ninguna	
Secuencia normal	Pasos	Acción
	1	Almacenar datos del usuario
	2	Almacenar puntajes
Postcondición	Ninguna	
Frecuencia	Varias ocasiones en la semana	
Comentarios	Los datos solo se almacenarán de forma local	

Diagrama de secuencia

Describe la secuencia que sigue el usuario con respecto de la aplicación. En la sección uno (serial válido) se realiza la validación del serial insertado por el usuario. En la sección dos (datos válidos) se validan los datos del usuario con los que hay en la base de datos, si existen esos datos, no le permitirá realizar el registro, sino el usuario quedará registrado. En la sección 3 el sistema valida que los datos dados por el usuario sean los mismo que hay en la base de datos. En la sección 4 el usuario entra en alguno de los niveles 1 o 2 el sistema evalúa la pronunciación, y la va guardando en la base de datos. La sección 5 se crea una sesión con Watson Assistant el cual enviará un saludo en texto al servicio de Text to Speech, el cual realizará un audio del texto y lo enviará al sistema quien reproducirá el audio al usuario. El usuario responderá, el sistema grabará lo que ha dicho el usuario y lo pasará al servicio de Speech to Text, quien generará un texto digital a partir de lo que el usuario dijo; Listo el texto digital se transferirá a Watson Assistant, con la idea de interpretar lo dicho por el usuario, después de esto enviará una respuesta al servicio de Text to Speech, quien lo retornará al usuario. (Diagrama 9 Diagrama de secuencia del usuario)

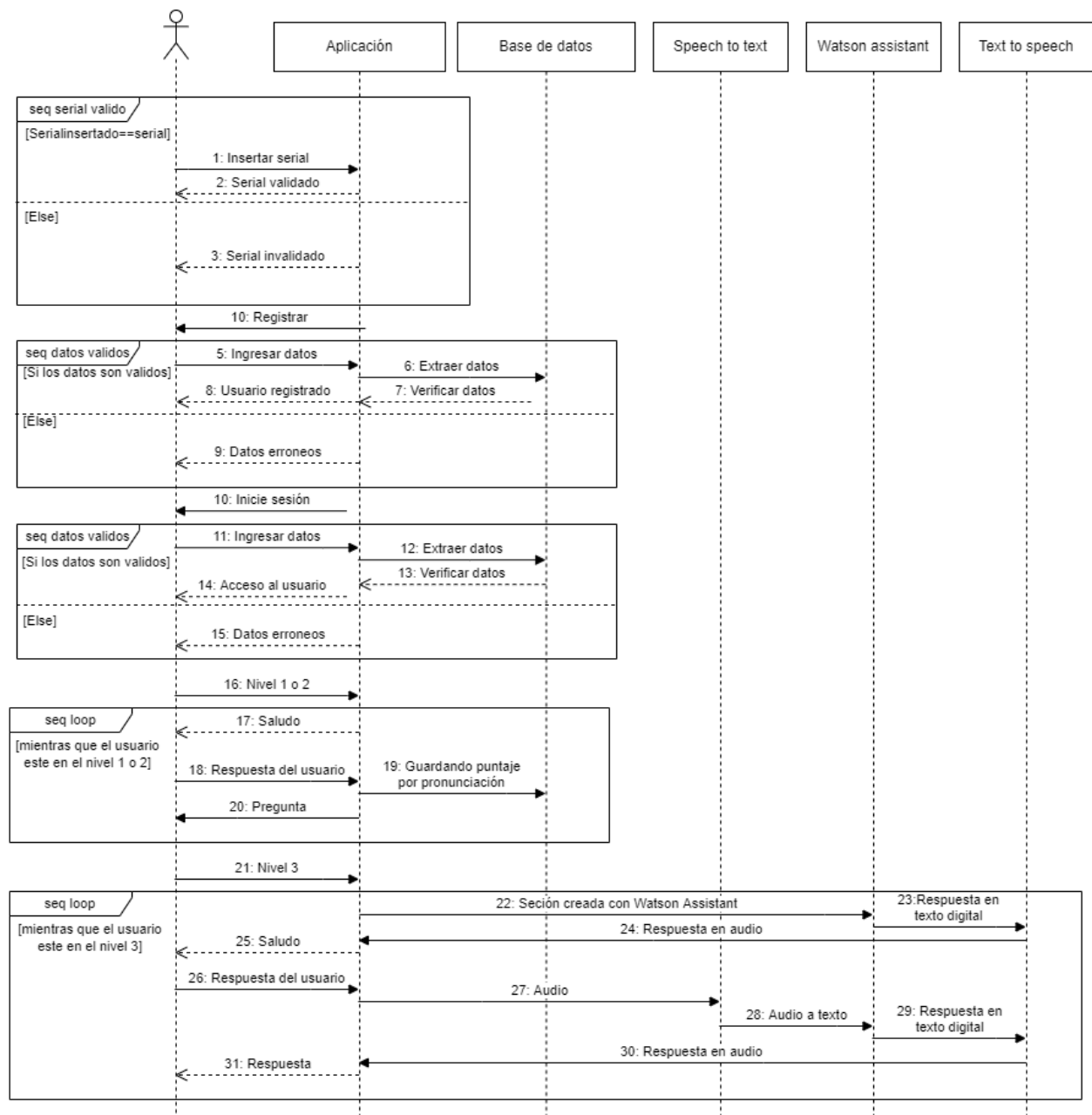


Diagrama 9 Diagrama de secuencia del usuario

Diagrama de despliegue

Este diagrama describe los componentes necesarios para que la aplicación funcione de una forma apropiada. Hay a tener en cuenta que la aplicación necesita internet para acceder a las funcionalidades relacionadas con las tecnologías (Watson, TTS, STT) que están en la nube de IBM Cloud. El sistema es desarrollado en un modelo para Windows 10 o superiores, y que, para hablar, el usuario debe configurar en el sistema, el micrófono que va a utilizar como por defecto, así también el parlante por donde va a recibir las repuestas del asistente.

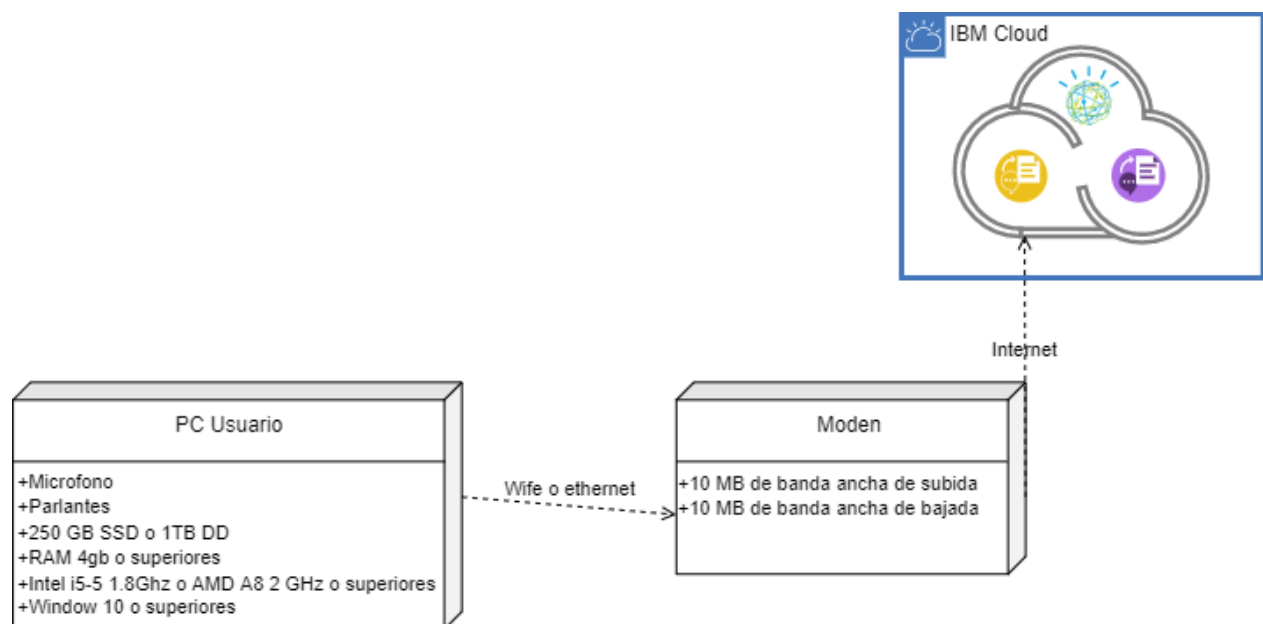


Diagrama 10 Diagrama de despliegue

Aplicación

Como solución a la ausencia de oportunidades de práctica de conversación que pueden presentar los estudiantes de este idioma, se plantea el desarrollo de sistema que integra el asistente virtual de IBM Watson junto con las tecnologías de Text to Speech y Speech to Text, las cuales permitirán que los estudiantes practiquen inglés conversacional.

Para ingresar al asistente el usuario deberá crear una cuenta local (los archivos no se guardarán en la nube), y de esta manera podrá ingresar al sistema a practicar conversación con el asistente, también se necesitará una conexión a internet, de lo contrario la aplicación será inservible.

De acuerdo con el nivel que el usuario haya seleccionado, podrá así mismo acceder a los niveles siempre y cuando sean menor o igual al que seleccionó. El asistente cuenta con dos niveles de inglés para práctica de pronunciación y el último para realizar la conversación.

En caso de que el usuario no sepa que hacer dentro del sistema hay una sección de ayuda la cual son una serie de imágenes las cuales explican el uso del asistente. También el sistema tiene una sección de puntajes donde muestra un puntaje de 0 a 10 en la pronunciación que el usuario tiene, de acuerdo con el servicio Speech to Text.

Funcionamiento

En condiciones ideales el asistente comenzará en la sección del inicio de sesión (ilustración 8). Aquí es donde el usuario puede iniciar la sesión y registrarse si anteriormente no lo ha hecho. En caso de no tener una cuenta registrada en el dispositivo, deberá registrarse dando clic en el botón de Sign Up de allí lo tomará a la sección de registro.

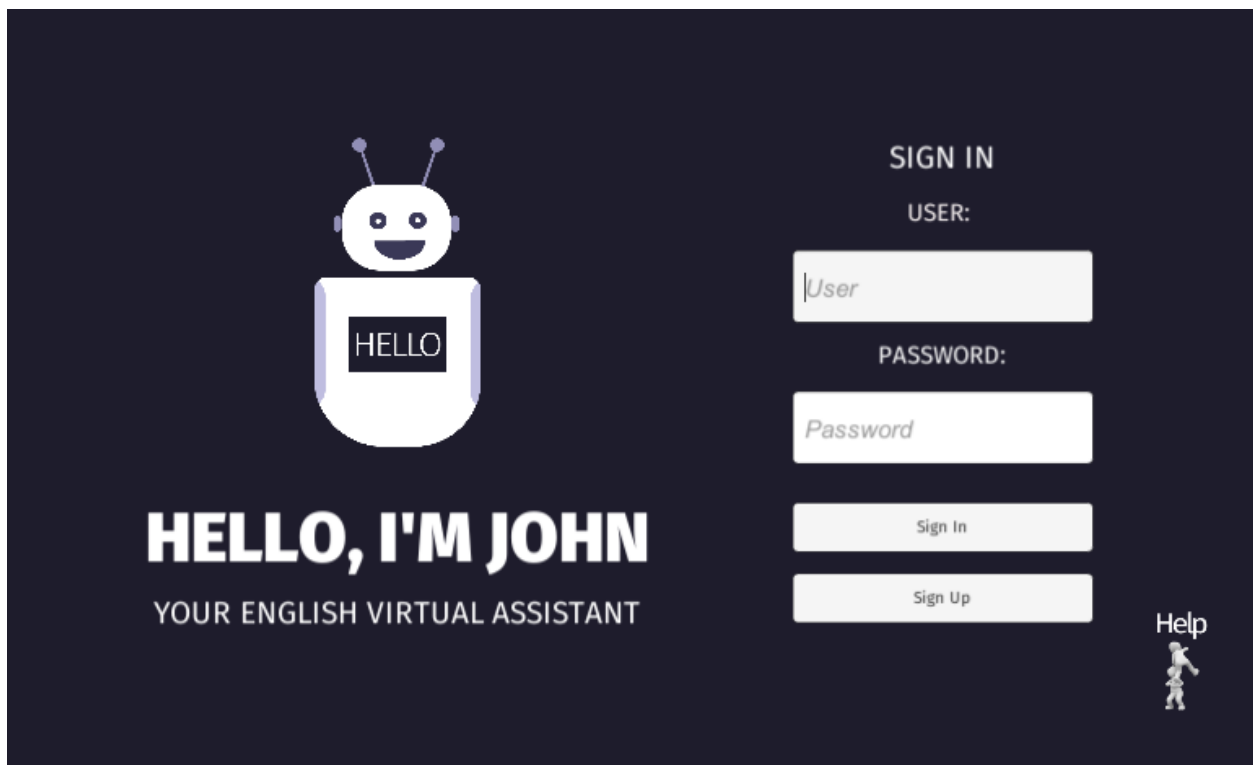
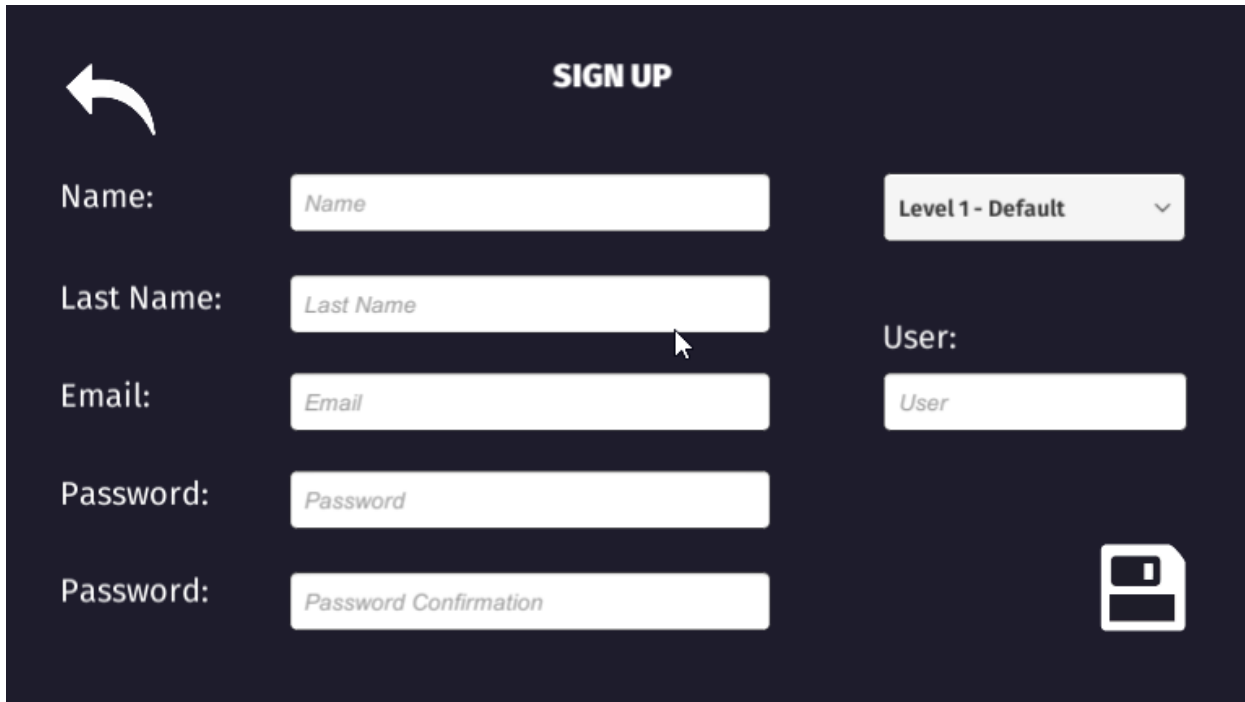


Ilustración 8 Inicio de sesión

Cuando el usuario haya dado clic en el botón Sign Up lo llevará al registro (ilustración 9). Aquí el usuario registrará los datos necesitados por la aplicación (nombre, apellido, email, contraseña, nivel de inglés, usuario), en caso de que un usuario ya exista con el mismo nombre o

correo, la aplicación avisará al usuario que se está registrando, que debe buscar un usuario distinto y un correo distinto, después de esto lo enviará a la sección de inicio de sesión para que inicie sesión.



The image shows a 'SIGN UP' form on a dark background. At the top left is a white curved arrow pointing left. The title 'SIGN UP' is centered at the top in white. The form contains the following fields and elements:

- Name:** A white input field with the placeholder text 'Name'.
- Last Name:** A white input field with the placeholder text 'Last Name'. A mouse cursor is pointing at the right side of this field.
- Email:** A white input field with the placeholder text 'Email'.
- Password:** A white input field with the placeholder text 'Password'.
- Password:** A white input field with the placeholder text 'Password Confirmation'.
- User:** A white input field with the placeholder text 'User'.
- Level:** A dropdown menu showing 'Level 1 - Default' with a downward arrow.
- Save Icon:** A white floppy disk icon located at the bottom right of the form.

Ilustración 9 Registro

Iniciada sesión, al usuario le aparecerá el menú donde tendrá varias opciones (Ilustración 10). Dentro de estas opciones aparecen los niveles, que si el usuario en sus configuraciones tiene nivel de inglés uno deberá cambiar de nivel para poder acceder al siguiente nivel, esto se puede realizar de dos maneras, la primera es practicando la pronunciación para alcanzar un puntaje mayor al 75% y pasando las 10 secciones del nivel 1, y la segunda opción es que el usuario se dirija a la opción de las configuraciones y manualmente cambie su nivel de inglés. Lo mismo pasa con el nivel 2, si ese fue seleccionado tendrá las mismas dos opciones, practicar

pronunciación o hacer el cambio manualmente el nivel en las configuraciones para pasar al nivel 3. Otras opciones que se pueden apreciar son los puntajes y el botón de ayuda, donde el usuario puede entrar y conocer el funcionamiento de la aplicación, hay que mencionar que el botón de ayuda, también está disponible al inicio de sesión.

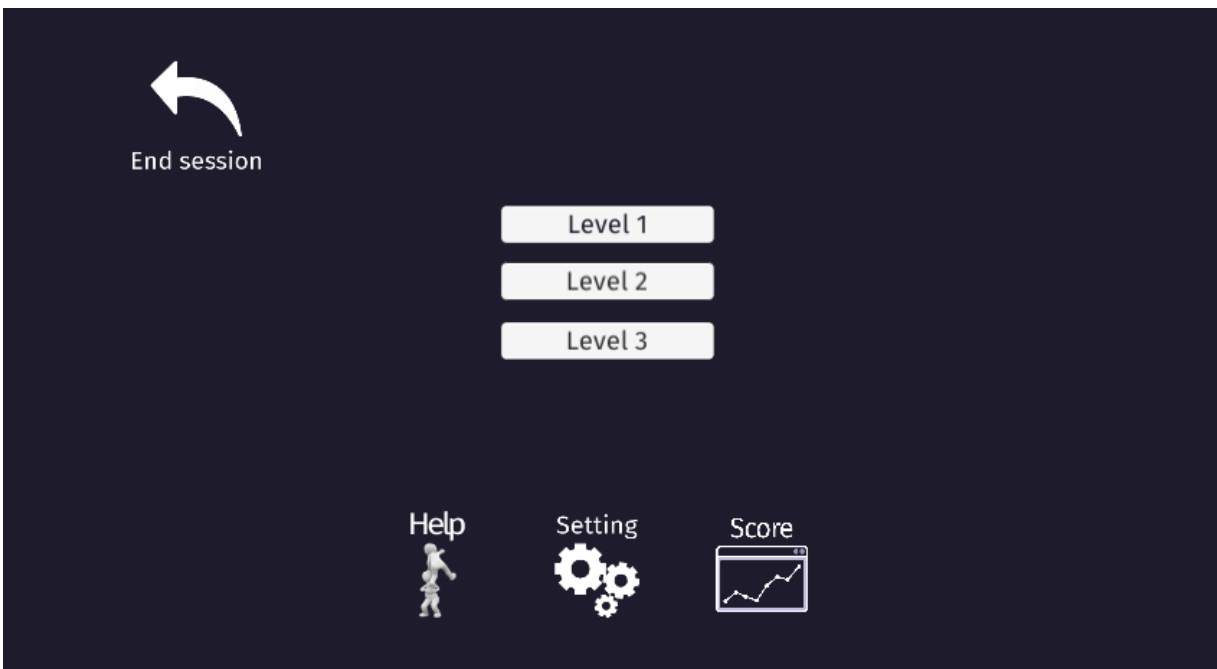


Ilustración 10 Menú

En el momento en que un usuario da clic en el botón ayuda, una nueva escena se mostrará. En esta nueva escena aparecerán una lista de opciones, con instrucciones de cómo funcionan las secciones que la aplicación contiene, por ejemplo: cómo registrarse, cómo funciona el menú, como funcionan las configuraciones, y cómo funcionan los niveles. Ilustración 11.

Hay que tener en mente que la aplicación es totalmente en inglés, así también, todas las instrucciones lo son.

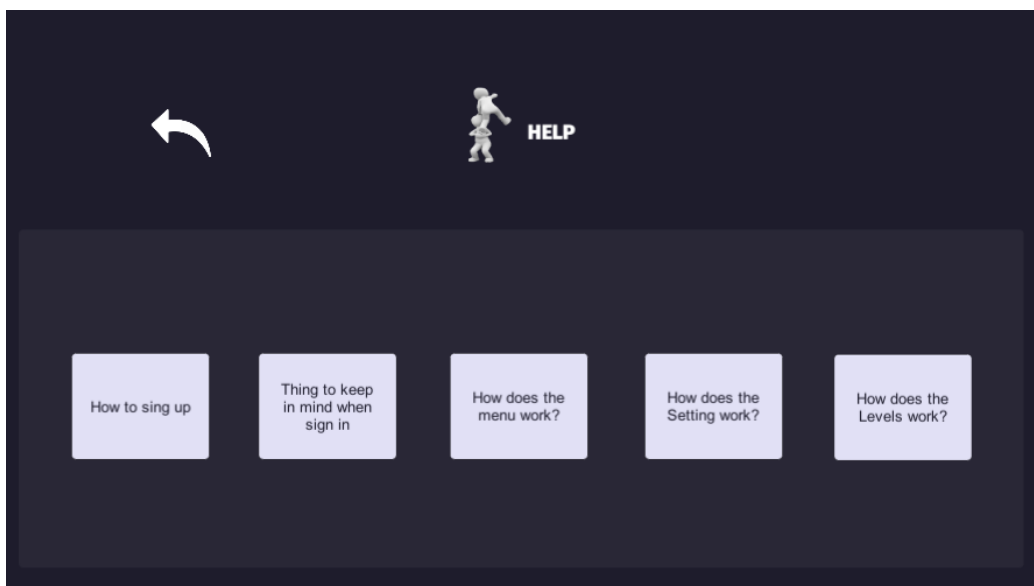


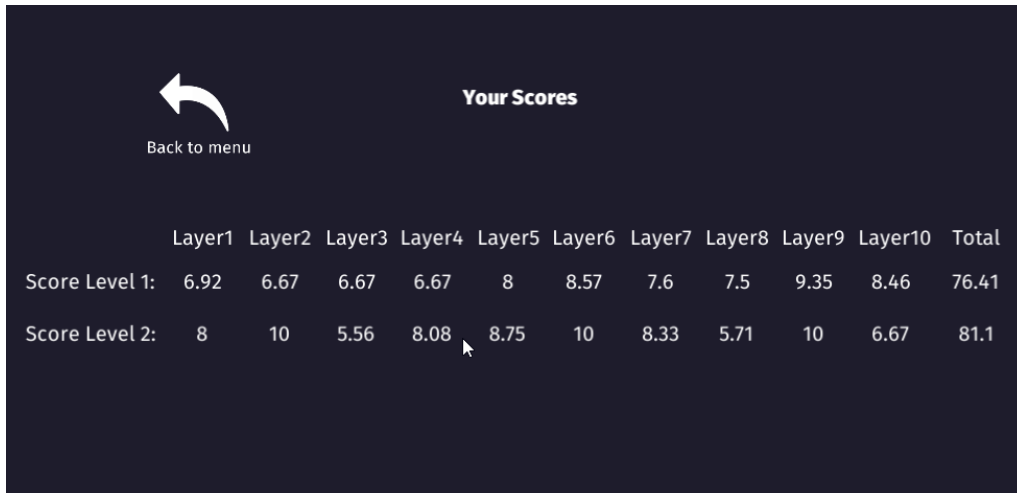
Ilustración 11 Ayuda

En el momento que el usuario quiera realizar modificaciones, puede entrar en la sección de las configuraciones (Ilustración 12) donde podrá realizar los cambios de nombre, apellido, contraseña, usuario etc.

A screenshot of a settings form interface. At the top, there is a white curved arrow pointing left, a gear icon, and the word "SETTINGS" in white capital letters. Below this, there are several input fields and a dropdown menu. On the left side, there are five rows of labels and input fields: "Name:" with a field containing "Name", "Last Name:" with a field containing "Last Name", "Email:" with a field containing "Email", "Old password:" with a field containing "Old password", and "New password:" with a field containing "New password". Below these, there is another "New password:" label with a field containing "Password Confirmation". On the right side, there is a dropdown menu labeled "Level 1 - Default" with a downward arrow, and a "User:" label with a field containing "User". At the bottom right, there is a white floppy disk icon.

Ilustración 12 Configuraciones

En la sección de los puntajes el sistema mostrará los puntos que el usuario ha ganado en los niveles uno y dos, calificando con 10 una pronunciación exacta o casi exacta, por cada layer o sección del nivel. Cada uno de estos dos niveles cuenta con 10 secciones. (Ilustración 13)



	Layer1	Layer2	Layer3	Layer4	Layer5	Layer6	Layer7	Layer8	Layer9	Layer10	Total
Score Level 1:	6.92	6.67	6.67	6.67	8	8.57	7.6	7.5	9.35	8.46	76.41
Score Level 2:	8	10	5.56	8.08	8.75	10	8.33	5.71	10	6.67	81.1

Ilustración 13 Puntajes

Cuando el usuario dentro del menú da clic en Level 1, eso lo llevará al primer nivel, donde comenzará a practicar la pronunciación. El sistema no dejara que avance hasta que haya oprimido el micrófono y haya leído la respuesta. Una vez hubo leído, obtendrá un puntaje.

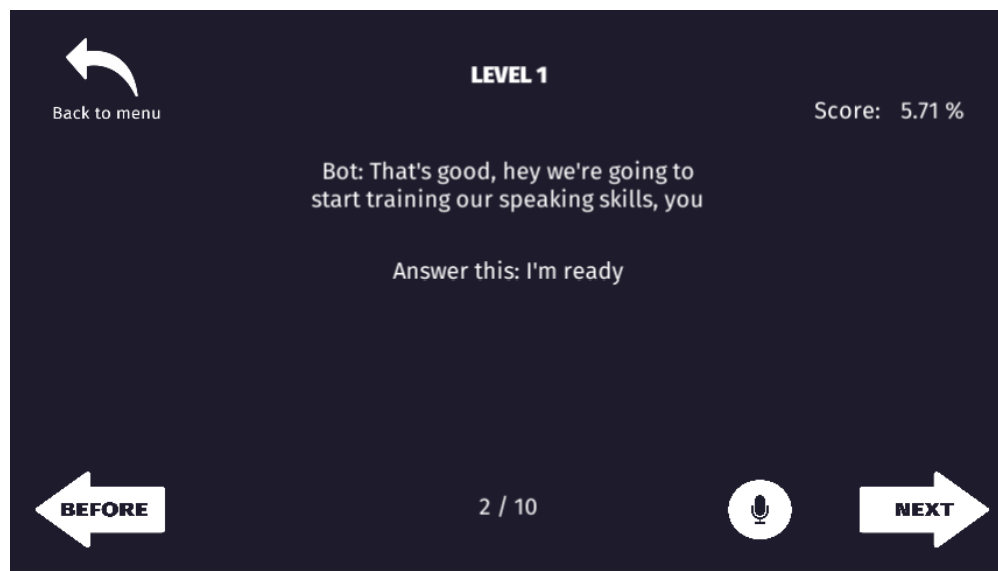


Ilustración 14 Nivel 1

En el nivel 2 (Ilustración 15), el usuario realiza la misma dinámica que utilizó en el nivel 1 (Ilustración 14).

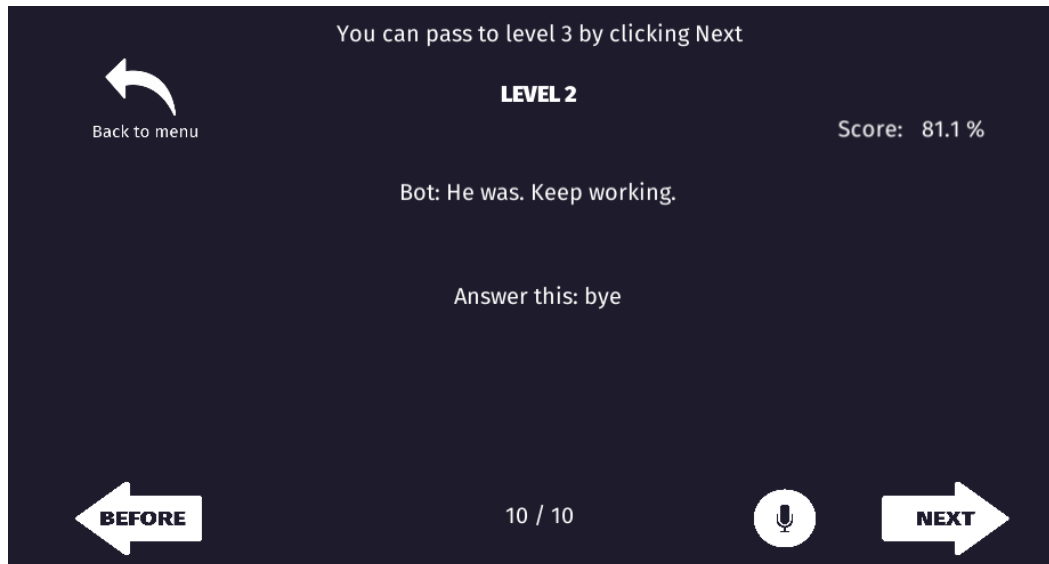


Ilustración 15 Nivel 2

En el nivel 3 (Ilustración 16) el usuario podrá interactuar con el asistente virtual de forma más dinámica, por lo cual el usuario no tendrá ningún texto que leer, sino que responderá de acuerdo al conocimiento que tenga en el momento que el asistente le pregunte sobre alguna cosa.

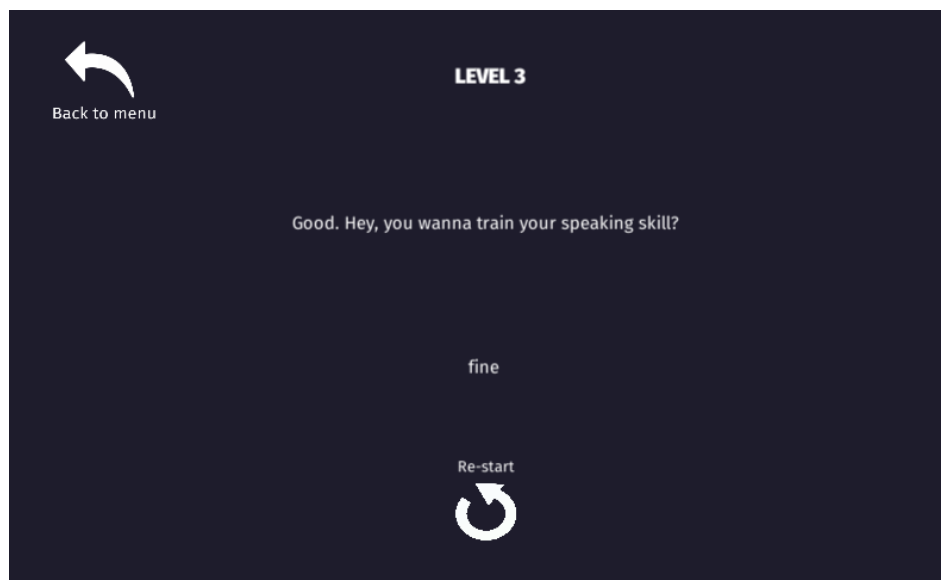


Ilustración 16 Nivel 3

Código

En las clases que envuelven los servicios se incluyeron las credenciales de acceso. En la Ilustración 17 Credenciales para conectar con IBM Watson muestra las credenciales que permiten conectar la aplicación con el servicio de Watson Assistant, estas credenciales autorizan a otras aplicaciones para poder utilizar los servicios de Watson, fuera del entorno de IBM.

```
public class AssistantBot : MonoBehaviour
{
    //----- Variables

    string assistantAPIkey = "0sIGwxE6sKqMXVtR3HR0U4z10LuQy1sQ69B1ITS61L ";
    string assistantURL = "https://api.au-syd.assistant.watson.cloud.ibm.com/instances/7f30a6df-6ae1-4f89-be9b-204156f2d945";
    string versionDate = "2020-09-24";
    string assistantID = "80094d8-8f52-4157-b8b9-3de0dcb9ae8";
}
```

Ilustración 17 Credenciales para conectar con IBM Watson

La Ilustración 18 Credenciales del servicio Speech to Text muestra las credenciales que conceden conectar la aplicación con el servicio de reconocimiento del habla.

```
public class SpeechToText : MonoBehaviour
{
    string sttAPIkey = "yyq4k7LEUdVPSBH5ZP8bc4brZuqbiWhlFsqjipNCg.";
    string sttURL = "https://api.au-syd.speech-to-text.watson.cloud.ibm.com/instances/7a3f0411-9cfe-4cae-84c8-b1c3245fa6b9";
}
```

Ilustración 18 Credenciales del servicio Speech to Text

La Ilustración 19 Credenciales del servicio Text to Speech muestra las credenciales que permiten conectar la aplicación con el servicio de texto a habla

```
public class TextToSpeech : MonoBehaviour
{
    string ttsAPIkey = "1HsHE1H6sfnFC162c2z8y-c_4hK79vH921p_RgnheK ";
    string ttsURL = "https://api.au-syd.text-to-speech.watson.cloud.ibm.com/instances/1539056a-7b38-486c-869f-89799ed757f5";
}
```

Ilustración 19 Credenciales del servicio Text to Speech

Para poder usar estos servicios es necesario crearlos, con este método se ayudan a crear el servicio de Text to Speech. Ilustración 20

```
private IEnumerator CreateService()
{
    //Debug.Log("TTS CreateService");//being used
    tts_authenticator = new IamAuthenticator(apikey: ttsAPIkey);

    // Wait for tokendata
    while (!tts_authenticator.CanAuthenticate())
    {
        yield return null;
    }
    //-----
    TTSService = new TextToSpeechService(tts_authenticator);
    TTSService.SetServiceUrl(ttsURL);
    //-----
}
```

Ilustración 20 Creación del servicio TTS

También es necesario crear los servicios de Speech to Text (Ilustración 21) para que el asistente pueda recibir información del usuario.

```
private IEnumerator CreateService()
{
    IamAuthenticator authenticator = new IamAuthenticator(apikey: sttAPIkey);

    while (!authenticator.CanAuthenticate())
    {
        yield return null;
    }
    _service = new SpeechToTextService(authenticator);
    _service.SetServiceUrl(sttURL);
    _service.StreamMultipart = true;

    Active = true;
}
```

Ilustración 21 Servicio STT

Así mismo para poder usar el asistente desde la aplicación, se necesita crear este servicio, con el método CreateService (Ilustración 22) se realiza este paso.

```
public IEnumerator CreateService()
{
    var authenticator = new IamAuthenticator(apikey: assistantAPIkey);
    while (!authenticator.CanAuthenticate())
    {
        yield return null;
    }
    assistantService = new AssistantService(versionDate, authenticator);
    assistantService.SetServiceUrl(assistantURL);
    assistantService.CreateSession(OnCreateSession, assistantID);
    while (!createSessionTested)
    {
        yield return null;
    }
}
```

Ilustración 22 Creación del servicio del asistente IBM Watson

IBM Watson Assistant

Intenciones:

El asistente está orientado para una conversación sencilla, con dos tópicos los cuales son fútbol y carros, todas las intenciones (Ilustración 23 y Tabla 11) creadas tiene una orientación a alguno de estos dos temas de conversación.

☐ Intents (47) ↑	Description
☐ #Bad	Expressions that user says because something doesn't see good
☐ #Bad_player	The user express that he is a bad player
☐ #Best_player	
☐ #Confirmation	
☐ #Conversation	
☐ #Do_not_know	
☐ #Do_not_like	
☐ #Do_not_play_often	
☐ #Do_not_remember	
☐ #Do_not_understand	When the person does not understand the question or needs to listen the questio...
☐ #Dont_like_game	
☐ #Favorite	
☐ #Feel_normal	The user doesn't enthusiastic or disappointed about something
☐ #FeelBad	
☐ #FeelGood	
☐ #General_About_You	Request generic personal attributes.
☐ #General_Ending	End the conversation.

Ilustración 23 Intenciones en la plataforma IBM Cloud

Fuente: <https://au-syd.assistant.watson.cloud.ibm.com/>

Como se pudo apreciar en el capítulo 3 de este documento, las intenciones utilizadas están relacionadas con expresiones que pueden ser vistas dentro de los contextos propuestos. La siguiente tabla muestra todas las intenciones creadas en el asistente Watson. (tabla 11)

Bad	Do_not_know	Why_I_like_road_trip
Bad_player	Do_not_like	Why_like_loud_car
FeelBad	Do_not_play_often	General_About_You
Best_player	Do_not_remember	General_Ending
FeelGood	Do_not_understand	General_Human_or_Bot
Good	Dont_like_game	General_Jokes
Good_player	I_dont_have	General_Positive_Feedback
How_often_play_football	Negation	So_so
Problem_on_road	Never_played	Feel_normal
Knowing	Favorite	GiveThanks
Knowledge	I_like_it	Greeting
WhatYouKnow	Like_Drive	Conversation
Confirmation	Like_Football	JokeConversation
I_have_played	Like_Run	Whatever
I_think	Like_slow	Laugh

Tabla 11 Intenciones utilizadas para el asistente

Rojo		Amarillo		Azul oscuro		Blanco	
Pistacho		Naranja		Purpura			
Verde		Azul Cielo		Purpura oscuro			
Limon		Azul Salvaje		Gris			

Tabla 12 Colores guía

En la tabla 11 se muestran las intenciones por colores, donde el color rojo representa una situación mala, dentro de la intención del usuario, mientras que en la de color pistacho una intención buena. El color verde concentrado representa la intención cuando el usuario está

intentando describir alguna situación, y el limon representa el interés del usuario por saber que conoce el asistente.

Las intenciones en amarillo, son aquellas relacionadas con las confirmaciones, cuando un usuario dice que, si ha jugado, la intención I_have_played es detectada en la expresión. En cambio, las de color naranja son aquellas detectadas cuando la expresión no es positiva. Las expresiones en azul cielo claro representan otro tipo de expresiones positivas relacionadas con gustos del usuario.

Las intenciones representadas en el azul salvaje son generalidades que se pueden presentar en una conversación, estas las trae por defecto el asistente de IBM. La intención en azul oscuro se representa aquellas frases donde no se habla ni bien ni mal de algo como ejemplo: “el partido estuvo más o menos”.

La intención en púrpura representa cuando el usuario no manifiesta ningún tipo de emociones en una sentencia, por ejemplo: “No es nada del otro mundo”

Las intenciones en morado oscuro son aquellas que están relacionadas con los saludos y las gracias.

La intención en gris es la que se activa cuando el usuario hace manifiesto que quiere tener una conversación con el asistente.

Las intenciones en blanco son aquellas que se encuentran en situaciones con bromas o risa, dentro de la categoría de intenciones no determinadas esta anything que se activa cuando el usuario no detecta o determina la intención dentro de lo dicho por el usuario.

Entidades:

Las entidades de este asistente están orientadas a nombres de países, jugadores y equipos. También hay años, colores, tamaños o medidas, unidades de medida, carros, tipos de carros, verbos, etc.

☐	Entity (21) ↑
☐	@Adjectives
☐	@Adverb
☐	@Automobile
☐	@Colors
☐	@Continents
☐	@Countries
☐	@English
☐	@Football_Award
☐	@Footballer
☐	@Football_position
☐	@Kind_of_cars
☐	@Nouns
☐	@Quantity

Ilustración 24 Entidades en la plataforma IBM Cloud

Fuente: <https://au-syd.assistant.watson.cloud.ibm.com/>

En la Table 13 se presenta un listado de entidades utilizadas por el asistente, estas entidades guardan algún tipo de relación con los tópicos utilizados

Adjectives	Nouns
Adverb	Quantity
Automobile	Size
Colors	Sport
Continents	Teams
Countries	Transportation
English	Unit
Football_Award	Verbs
Footballer	World_cup
Football_position	Years
Kind_of_cars	

Tabla 13 Entidades utilizadas por el asistente

- Adjectives: Cuando se mencionan adjetivos, ej: Ruido, Clásico.
- Adverb: Cuando se mencionan adverbios, ej: Probablemente.
- Automobile: Cuando se mencionan nombres de automóvil, ej: Land Rover, Kia.
- Colors: Cuando se mencionan colores, ej: Blue, Red, Yellow.
- Continents: Cuando se mencionan continentes, ej: Europe, Asia.
- Countries: Cuando se mencionan países, ej: Colombia, England
- English: Cuando se menciona el idioma inglés, ej: English
- Football_Award: Cuando se mencionan premios award, ej: Best Coach of the year
- Footballer: Cuando se mencionan jugadores, ej: Jack Grealish, Freddy Rincón
- Football_position: Cuando se mencionan posiciones de juego, ej: Deep-Lying Playmaker
- Kind_of_cars: Cuando se mencionan tipos de carro, ej: Muscle car, Limousine
- Nouns: Cuando se mencionan sustantivos, ej: Trip

- Quantity: Cuando se mencionan cantidades, ej: Much, many
- Size: Cuando se mencionan tamaños, ej: Far, Long, Big
- Sport: Cuando se mencionan deportes, ej: Football, American Football, Soccer
- Teams: Cuando se mencionan equipos de football, ej: Valencia, Colombian team
- Transportation: Cuando se mencionan medios de transportes, ej: Motorcycle, Plane
- Unit: Cuando se mencionan unidades de medida metros o millas, ej: Meter, Mile
- Verbs: Cuando se mencionan verbos, ej: Like, Feel, Love, Run
- World_cup: Cuando se mencionan copas del mundo, ej: FIFA world cup 1974
- Years: Cuando se mencionan años, ej: 2022, 2020

Diálogos:

Los diálogos están en función de las dos líneas de conversación que el asistente tiene, y de acuerdo a las condiciones que tiene programadas, va dirigir el flujo de la conversación.

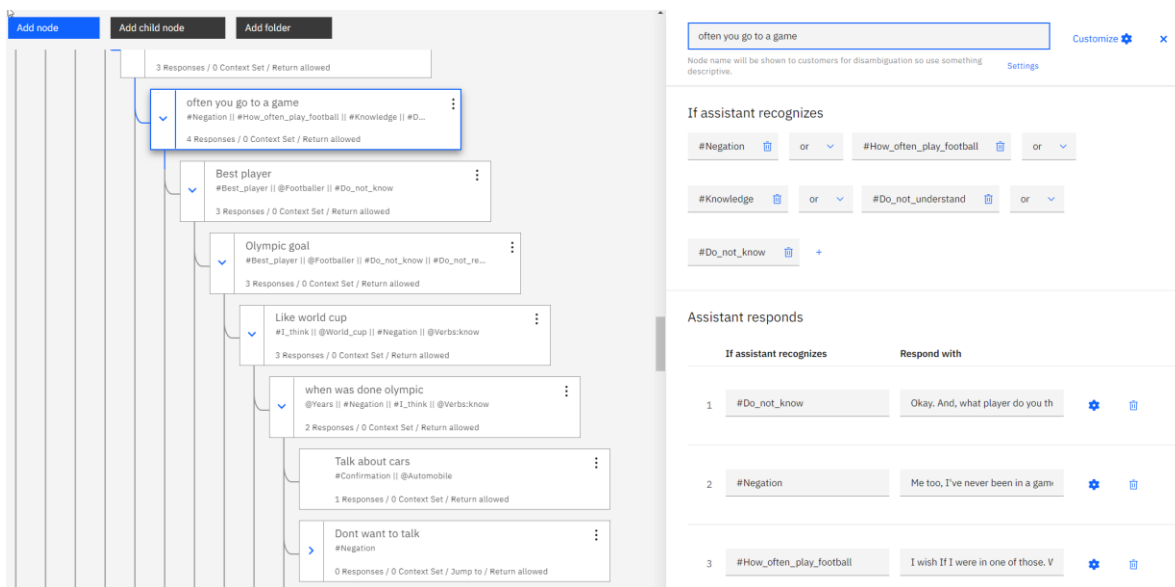


Ilustración 25 Diálogos en la plataforma IBM Cloud

Fuente: Propio de la cuenta en IBM Cloud

Como funcionan los diálogos:

El asistente inicia con un saludo en el nodo Welcome:

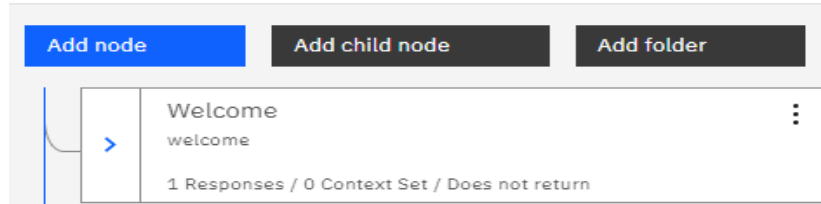


Ilustración 26 Nodo Welcome

Fuente: Propio de la cuenta en IBM Cloud

Donde tiene unos saludos que iniciarán la conversación, esto aparecerán de forma aleatoria.

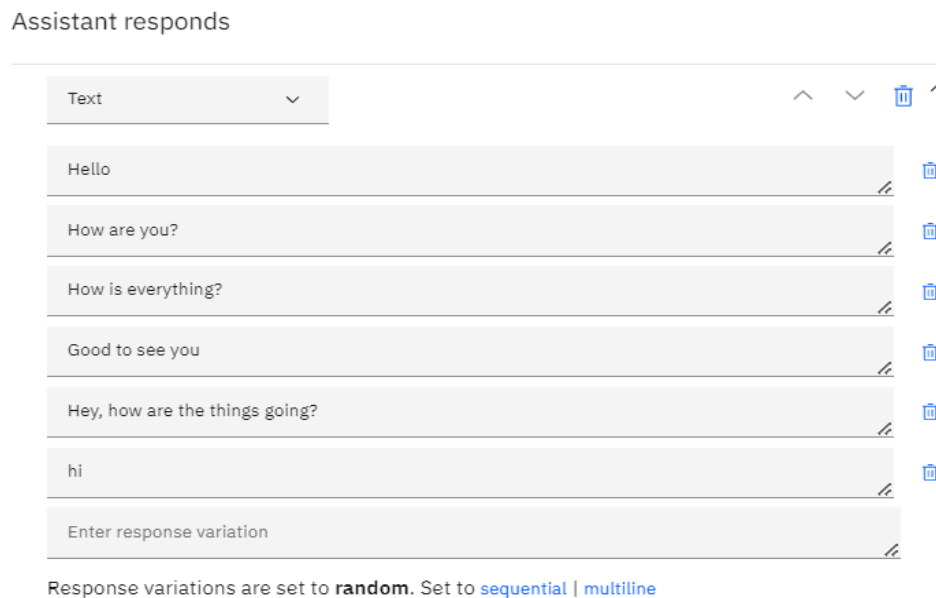


Ilustración 27 Variedad de saludos en el nodo Welcome

Fuente: Propio de la cuenta en IBM Cloud

Para este ejemplo se seleccionará el saludo: “Hey, how are the things going?” a lo cual el usuario puede responder: “fine” después de esto el asistente evaluará la respuesta, en este caso entrará en el nodo hijo del nodo Welcome que es How are thing, Ilustración 28.



Ilustración 28 Nodo How are thing

Fuente: Propio de la cuenta en IBM Cloud

Porque la expresión “fine” satisface las intenciones de #Confirmation o #Good, después el asistente dará una respuesta de las que él tiene allí. Se tomará como ejemplo: “Good. hey, you wanna train your speaking skill?” como se muestra en la ilustración 29.

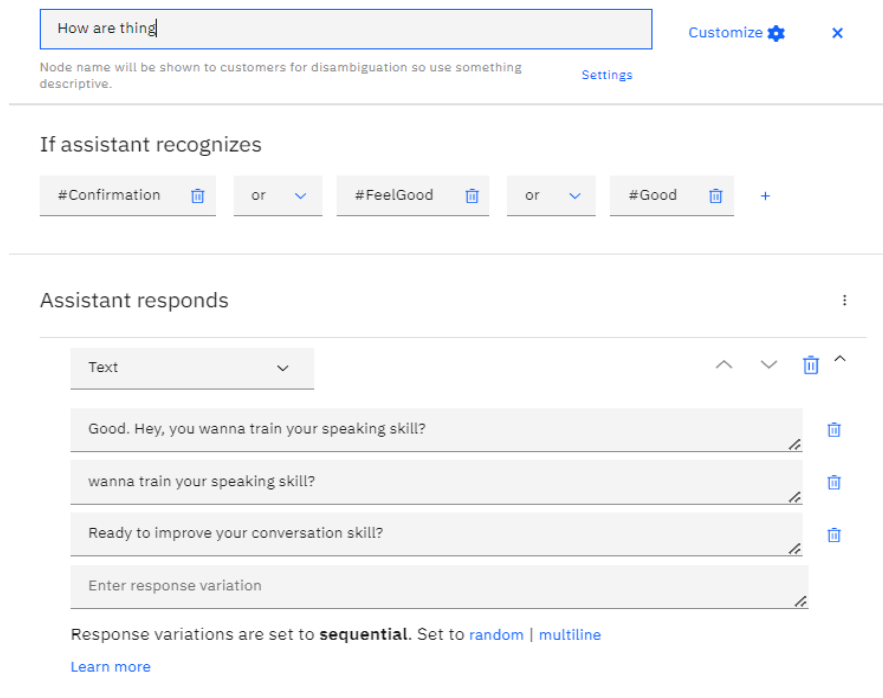


Ilustración 29 Condiciones y respuestas del nodo How are thing

Fuente: Propio de la cuenta en IBM Cloud

Para esto el usuario puede responder “yeah, I want to train” lo cual entrará en otro nodo que contenga intenciones que sean satisfechas con la respuesta dada por el usuario, como el About What Do You Want to Talk. Ilustración 30.

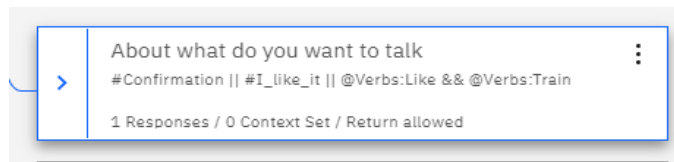


Ilustración 30 Nodo About What Do You Want to Talk

Fuente: Propio de la cuenta en IBM Cloud

En este caso entraría por la confirmación que está dentro de la sentencia “yeah, I want to train” y el asistente responderá “I know Little about cars and football. Choose one of those” y esperará la respuesta del usuario.

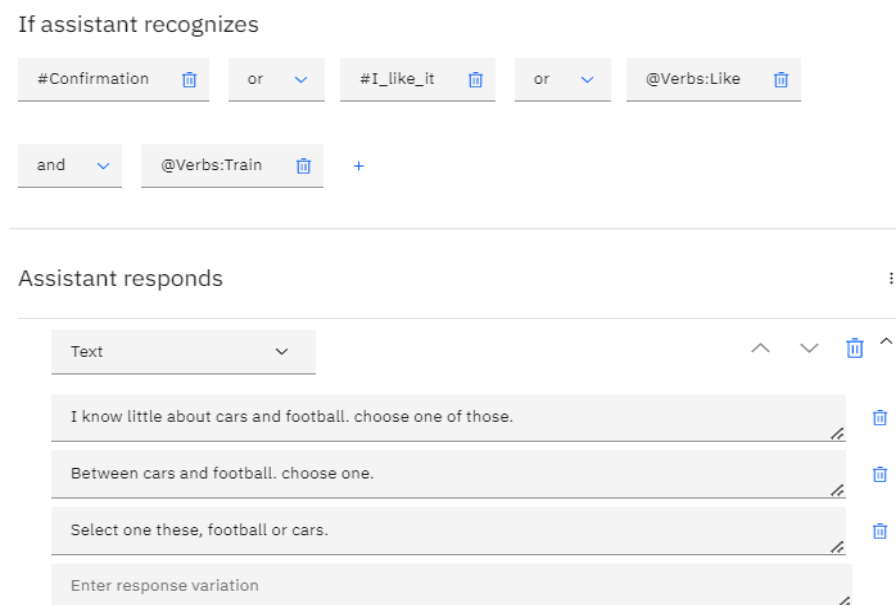


Ilustración 31 Condiciones y respuestas del nodo About What Do You Want to Talk

Fuente: Propio de la cuenta en IBM Cloud

En este punto el usuario debe de tomar una decisión, la cual podría ser “football” o “car”, para este ejemplo se seleccionará “football”, lo que lo pasará al nodo de Conversation y consecuente al nodo hijo About football (Ilustración 33), el asistente le dirá al usuario que: “I know about football, not to much. You like it?” en caso de no haber una respuesta concisa del usuario antes de pasar por este nodo, el asistente podría responder con: “car or football?” como se aprecia en la Ilustración 32, en la parte 3 de lo que el asistente reconoce.

Conversation

Customize

Node name will be shown to customers for disambiguation so use something descriptive. [Settings](#)

If assistant recognizes

#Conversation +

Assistant responds

	If assistant recognizes	Respond with	Finally		
1	@Sport:Football @Sport:Soc	Enter a response	Jump		
2	@Automobile:car	Enter a response	Jump		
3	true	car or football?	-		

Ilustración 32 Condiciones y respuestas del nodo Conversation

Fuente: Propio de la cuenta en IBM Cloud

Si el usuario respondió de forma concisa, el asistente lo llevara al nodo correspondiente, en este caso es About football. Ilustración 33.

The screenshot displays the configuration interface for a node named 'About football'. At the top, there is a text input field containing 'About football', followed by 'Customize' and 'Settings' links. Below this, a note states: 'Node name will be shown to customers for disambiguation so use something descriptive.' The main section is titled 'If assistant recognizes' and contains a sequence of conditions: '@Sport', 'or', '#Like_Football', 'or', and '#Whatever'. Each condition is in a box with a trash icon. Below this is the 'Assistant responds' section, which shows a list of possible responses under a 'Text' dropdown. The responses are: 'I know about Football, not much. You like it?', 'I've learned about that; tell me, you like football?', and 'You like football?'. Each response has a trash icon and a small icon indicating it can be expanded or collapsed.

Ilustración 33 Condiciones y respuestas del nodo About football

Fuente: Propio de la cuenta en IBM Cloud

Aquí es donde la entidad @Sport esta como una de las condiciones, y como football entra en la categoría de deportes, la condición se puede cumplir. El usuario puede responder a la pregunta de: “I know about football, not to much. You like it?” con un: “Yes I do” o “I don’t like it”. Para este ejemplo se usará el “Yes I do”. La respuesta dada por el usuario será evaluada por el asistente, y como el nodo Like football (Ilustración 34) tiene condiciones que pueden ser satisfechas con la respuesta anterior, el asistente entrará en ese nodo.

Like football

Customize  

Node name will be shown to customers for disambiguation so use something descriptive. [Settings](#)

If assistant recognizes

#Like_Football 

or 

#I_like_it 

or 

#Confirmation 

or 

#So_so 

+

Assistant responds 

Text 






I like too. What's your favorite team?



I love it too. How often do you play football?



Fan of what team?



Enter response variation



Response variations are set to **random**. Set to [sequential](#) | [multiline](#)

[Learn more](#)

Ilustración 34 Condiciones y respuestas del nodo Like football

Fuente: Propio de la cuenta en IBM Cloud

Estando aquí el asistente preguntará al usuario alguna de las preguntas de forma aleatoria, para este ejemplo se usará “I like too. What’s your favorite team?” a lo que el usuario puede responder: “The Colombian team”. Finalmente, el usuario también puede despedirse del asistente usando palabras como: “Bye”. Esto llevará al usuario al nodo de despedida End_Conversation como se puede apreciar en la ilustración 35.

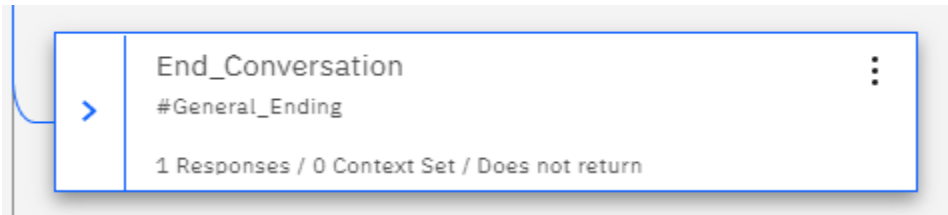


Ilustración 35 Nodo End_Conversation

Fuente: Propio de la cuenta en IBM Cloud

De este punto el asistente responderá con otro saludo de despedida. El usuario puede salir en cualquier momento de la aplicación.

Factores que pueden interrumpir la conversación

El asistente necesita un entrenamiento constante, lo que significa de que las conversaciones se pueden ver interrumpidas en las primeras etapas de su desarrollo, para esto la aplicación dispone de un botón “re-start” el cual permite reiniciar la conversación con el asistente. Estas interrupciones se pueden ocasionar cuando la intención del usuario no puede ser correctamente identificada, y ya que el asistente en la aplicación solo cuenta con entra y salida de texto, se ocasiona un bloqueo en el proceso, al no poder desplegar las opciones que él detecto como posibles intenciones para que el usuario pueda identificar la correcta intención que él tenía. Otro motivo por el cual se puedan presentar errores, es por la pronunciación de los usuarios, que a veces le resulta difícil al asistente entender.

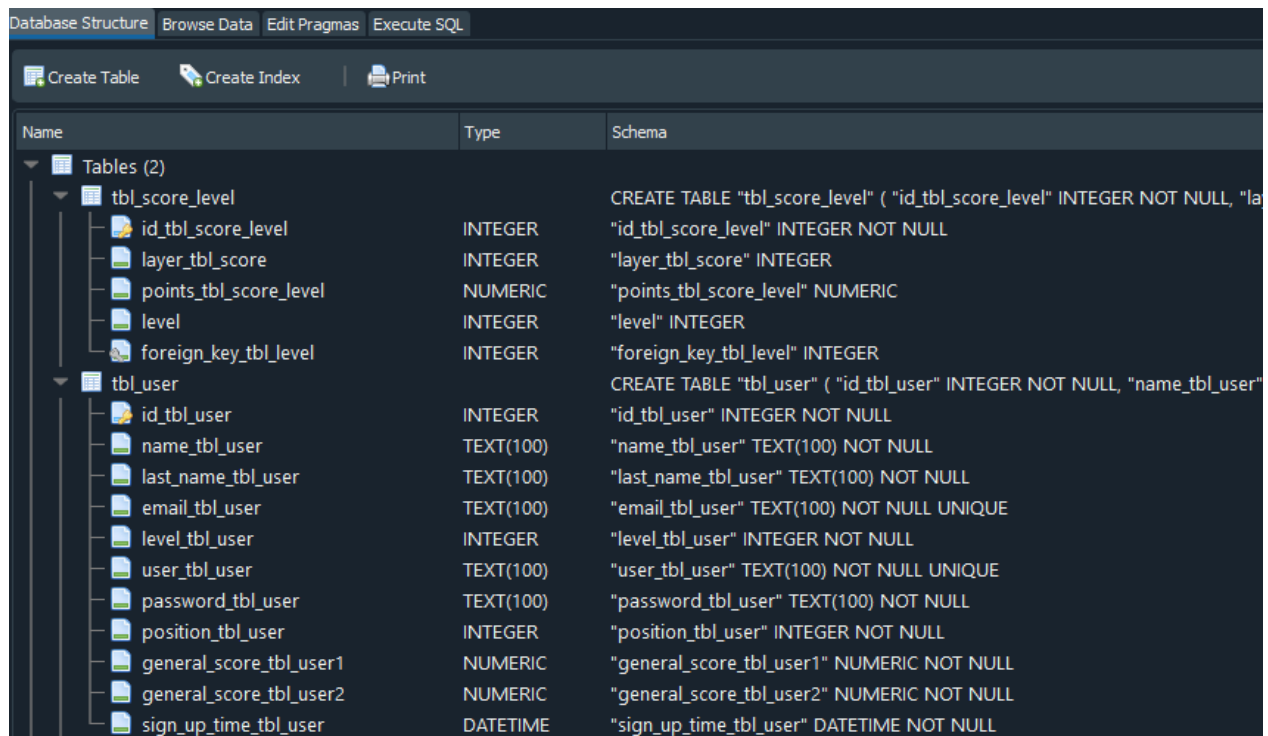
Conexión

La conexión entre los servicios de IBM Watson y la aplicación hecha en Unity con el SDK versión unity-sdk-5.0.2 que permite la creación de los servicios desde la aplicación, y el envío de información de ida y vuelta

Base de datos

El gestor de base de datos utilizado fue el de SQLite. Esta base de datos funciona de forma local, donde se guardarán los datos de los usuarios como contraseñas y niveles, también guardara los puntajes.

Cuando se inicia sesión o un usuario se registra, la base datos es llamada desde la aplicación, ya sea para guardar información, o validarla. Cuando la aplicación está corriendo en alguno de los niveles excepto el 3, los puntos se van guardando en la base datos de acuerdo al número o la posición dentro del nivel. Si un usuario obtuvo el puntaje requerido de 75 % para pasar al siguiente nivel, también se actualizar el nivel de dicho usuario dentro de la base de datos.



Name	Type	Schema
Tables (2)		
tbl_score_level		CREATE TABLE "tbl_score_level" ("id_tbl_score_level" INTEGER NOT NULL, "la
id_tbl_score_level	INTEGER	"id_tbl_score_level" INTEGER NOT NULL
layer_tbl_score	INTEGER	"layer_tbl_score" INTEGER
points_tbl_score_level	NUMERIC	"points_tbl_score_level" NUMERIC
level	INTEGER	"level" INTEGER
foreign_key_tbl_level	INTEGER	"foreign_key_tbl_level" INTEGER
tbl_user		CREATE TABLE "tbl_user" ("id_tbl_user" INTEGER NOT NULL, "name_tbl_user"
id_tbl_user	INTEGER	"id_tbl_user" INTEGER NOT NULL
name_tbl_user	TEXT(100)	"name_tbl_user" TEXT(100) NOT NULL
last_name_tbl_user	TEXT(100)	"last_name_tbl_user" TEXT(100) NOT NULL
email_tbl_user	TEXT(100)	"email_tbl_user" TEXT(100) NOT NULL UNIQUE
level_tbl_user	INTEGER	"level_tbl_user" INTEGER NOT NULL
user_tbl_user	TEXT(100)	"user_tbl_user" TEXT(100) NOT NULL UNIQUE
password_tbl_user	TEXT(100)	"password_tbl_user" TEXT(100) NOT NULL
position_tbl_user	INTEGER	"position_tbl_user" INTEGER NOT NULL
general_score_tbl_user1	NUMERIC	"general_score_tbl_user1" NUMERIC NOT NULL
general_score_tbl_user2	NUMERIC	"general_score_tbl_user2" NUMERIC NOT NULL
sign_up_time_tbl_user	DATETIME	"sign_up_time_tbl_user" DATETIME NOT NULL

Ilustración 36 Dos tablas mostradas en DB Browser for SQLite

Fuente: Propio, capture de pantalla de SQLite Browser

La base de datos relacional cuenta con dos tablas, una donde se almacena los datos de usuario, y la otra donde se almacenan los puntajes que el usuario realice en los niveles uno y dos. Esta última tabla es la que se usa más durante la ejecución de la aplicación.

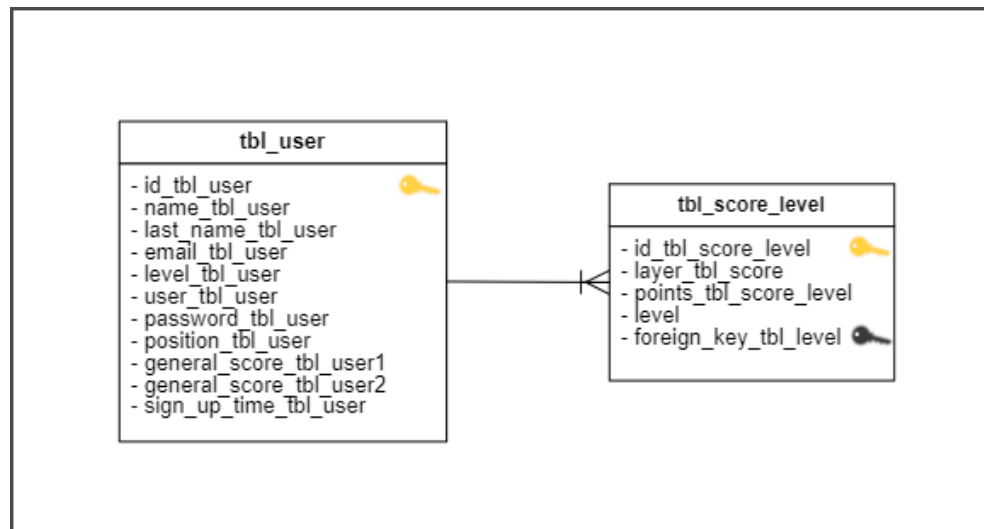


Ilustración 37 Base de datos relacional

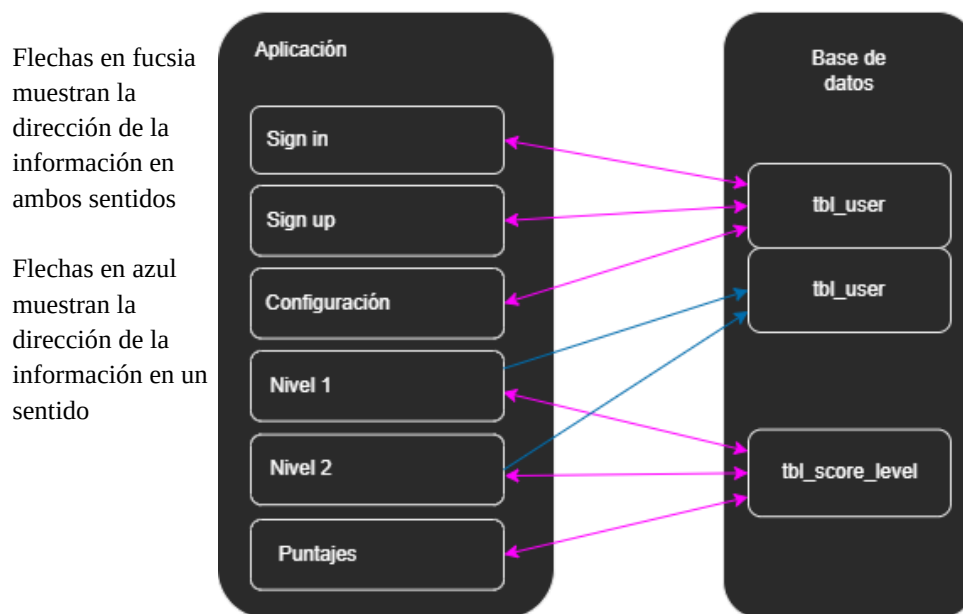


Ilustración 38 Interacción entre la base de datos y la aplicación

Cuando un usuario se registra, la aplicación hace una comprobación en la base de datos, esta comprobación busca que no se repitan los mismos datos en la tabla `tbl_user` que es la tabla de los usuarios, una vez un usuario se ha registrado puede acceder a la aplicación. Cuando el usuario accede a la aplicación, la misma hace una comprobación con los datos suministrados por el usuario, si los datos concuerdan con lo que hay en su sistema, importará los otros datos del usuario de la tabla `tbl_user`, y ubicará al usuario en el menú de la app. Con los datos importados realizará validaciones, que permitirán el acceso a los niveles.

Al momento en que el usuario va a ingresar a los niveles 1 o 2, estos piden a la base de datos, los puntajes que tengan registrado del usuario en la tabla `tbl_score_level`, que en caso de haberlos serán enviados a los respectivos niveles dentro de la aplicación. (Ilustración 38)

Si el usuario comienza a utilizar los niveles, estos comenzarán a interactuar con las dos tablas, actualizando la última posición del usuario en la aplicación y el puntaje general por nivel en la tabla `tbl_user`, y en la tabla `tbl_score_user` actualizando los puntajes individuales por sección.

Capítulo 5: Validación

Testeo y Datos Recolectados

El asistente se probó con estudiantes de Ingeniería de Sistemas que estaban cursando nivel I grupo AR y BR de inglés, añadido a esto también se probó con estudiantes del nivel III de grupo AR, en total fueron 11 personas las que probaron el asistente.

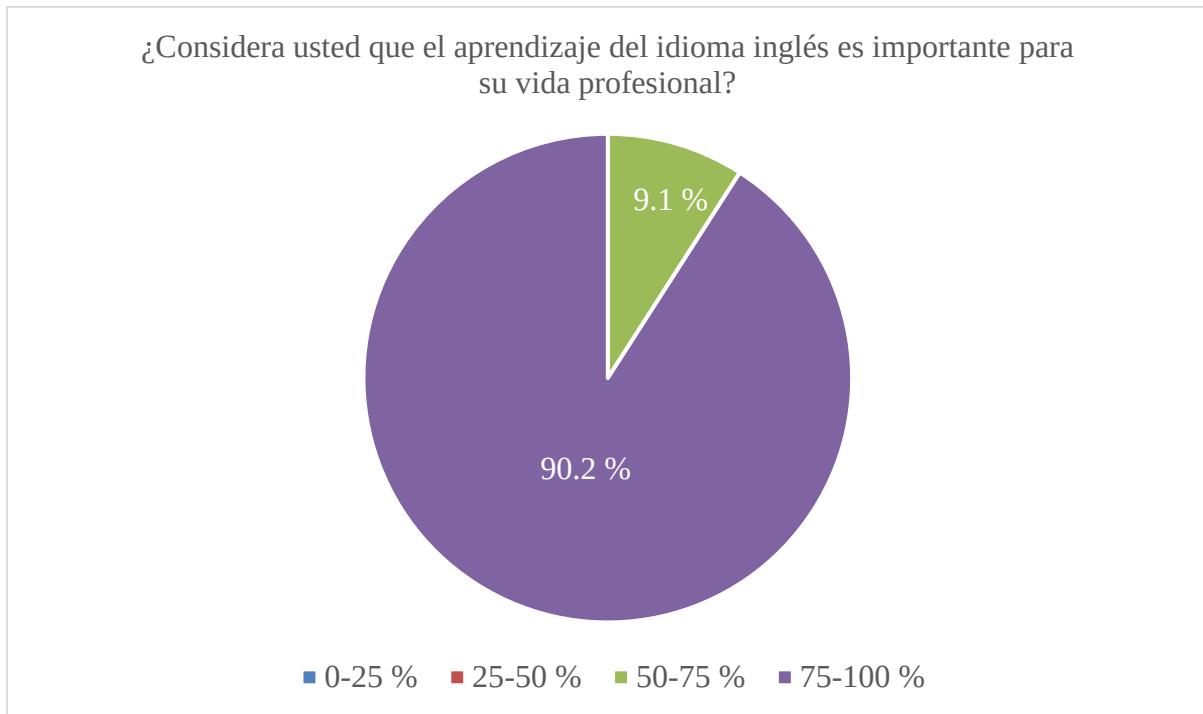
El 100 % de los estudiantes manifestaron ser de Ingeniería de Sistemas



Grafica 1 ¿A qué programa pertenece?

Fuente: Formularios de Google

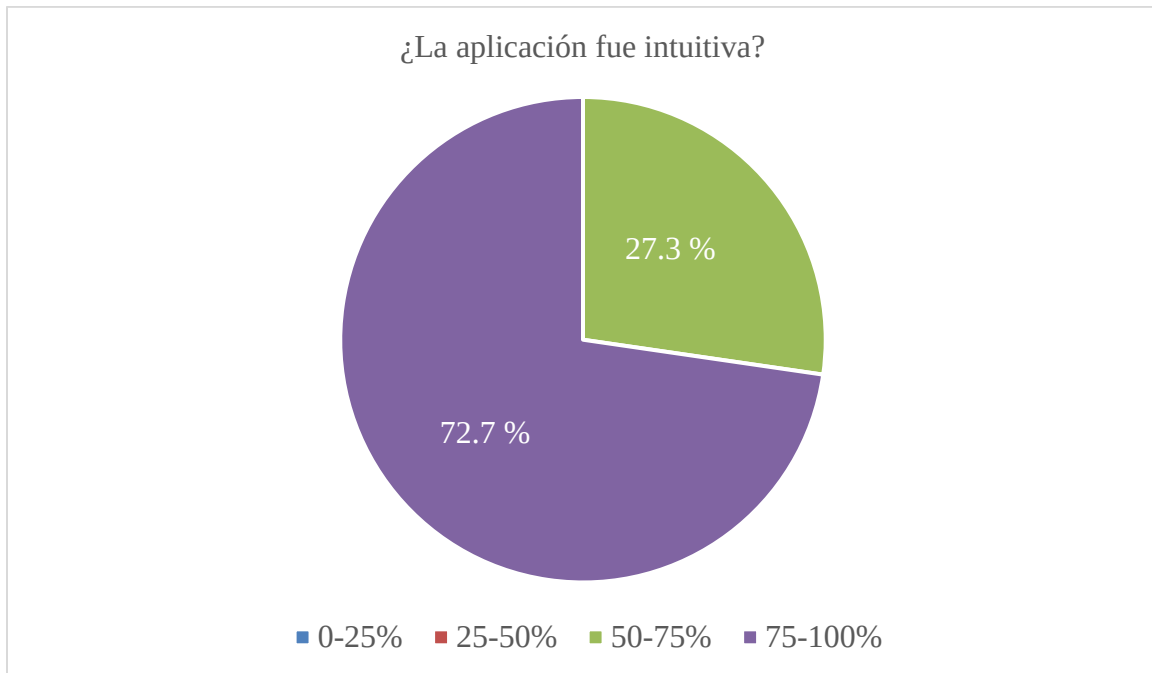
Un 90.9 % de los encuestados señaló que el inglés era importante entre un 75 a 100 % en su vida profesional, y un 9.1 % dijo que entre un 50 y un 75 %, como se muestra en la gráfica 2.



Grafica 2 ¿Considera usted que el aprendizaje del idioma inglés es importante para su vida profesional?

Fuente: Formularios de Google

Un total del 72.7 % dio a conocer que la aplicación fue intuitiva entre un 75 % y un 100 %, mientras un 27,3 % dijo que era entre un 50 y un 75 %, ver gráfica 3.

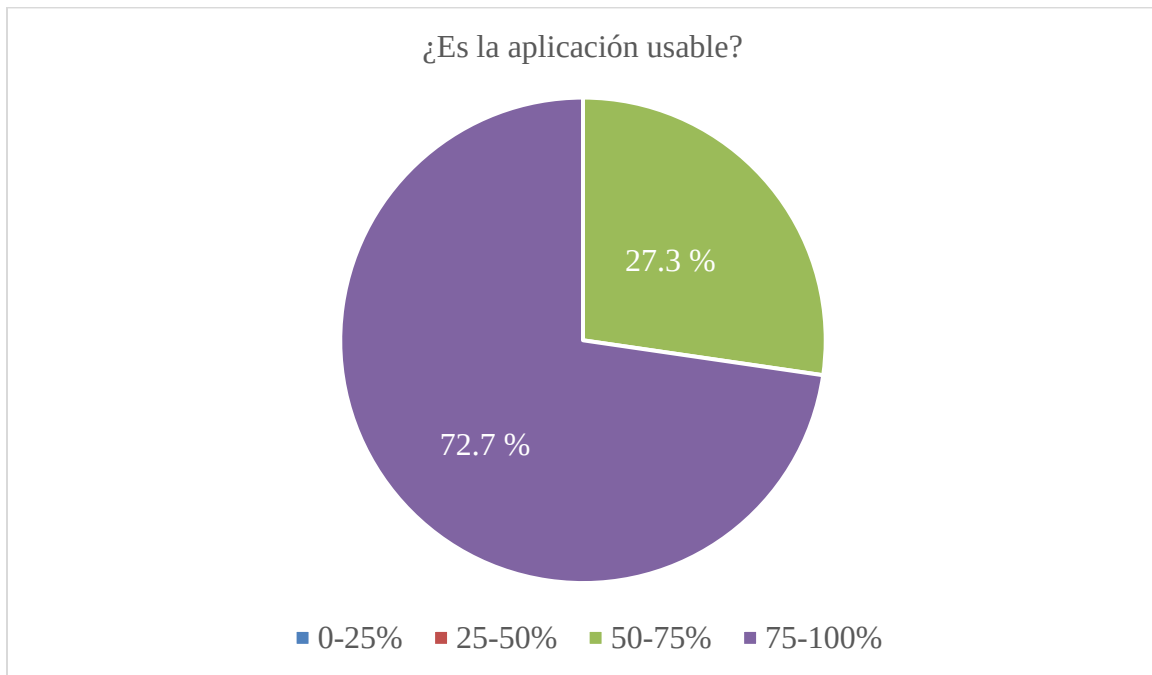


Gráfica 3 ¿La aplicación fue intuitiva?

Fuente: Formularios de Google

Esta información permite conocer que el total de los encuestados considera que la aplicación es intuitiva en más de un 50 %. Hay que tener en cuenta la aplicación está totalmente en inglés, lo que puede ser un poco complicado si no se entiende el idioma, pero también se incluyeron botón con imágenes, para que fuera más entendible.

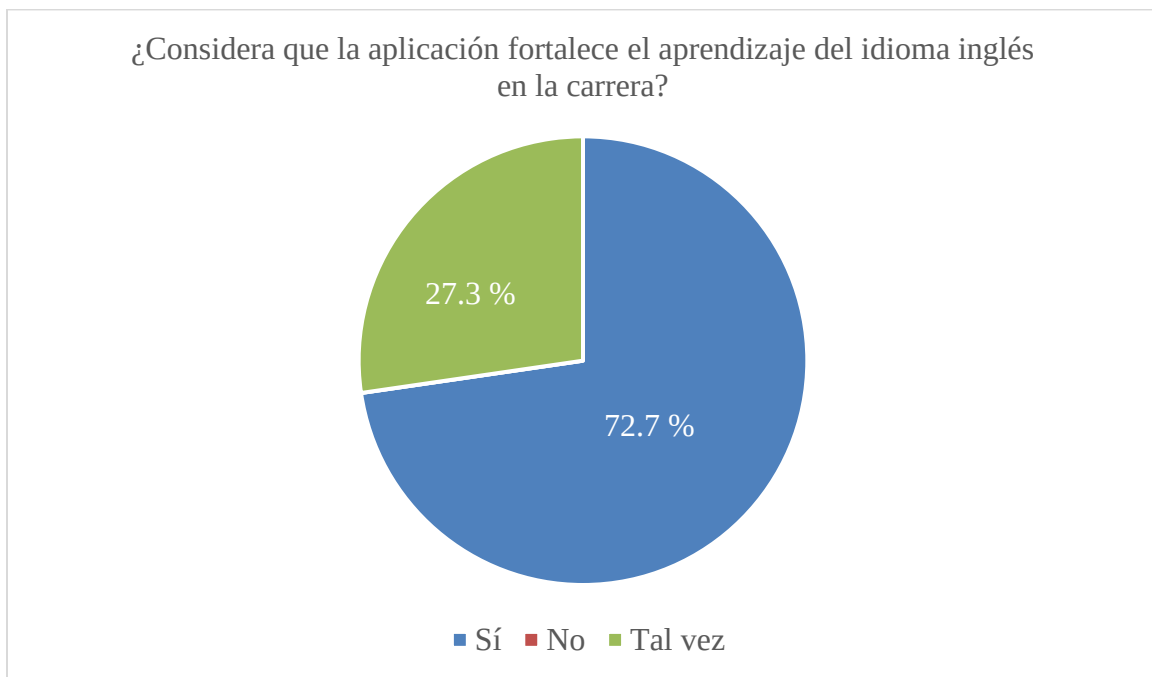
En la usabilidad de la aplicación un 72.7 % mencionó que la aplicación era usable entre el 75 % y el 100 %, a diferencia del 27.3 % de los encuestados dio a conocer que lo era entre un 50 y un 75 %, observar gráfica 4.



Grafica 4 ¿Es la aplicación usable?

Fuente: Formularios de Google

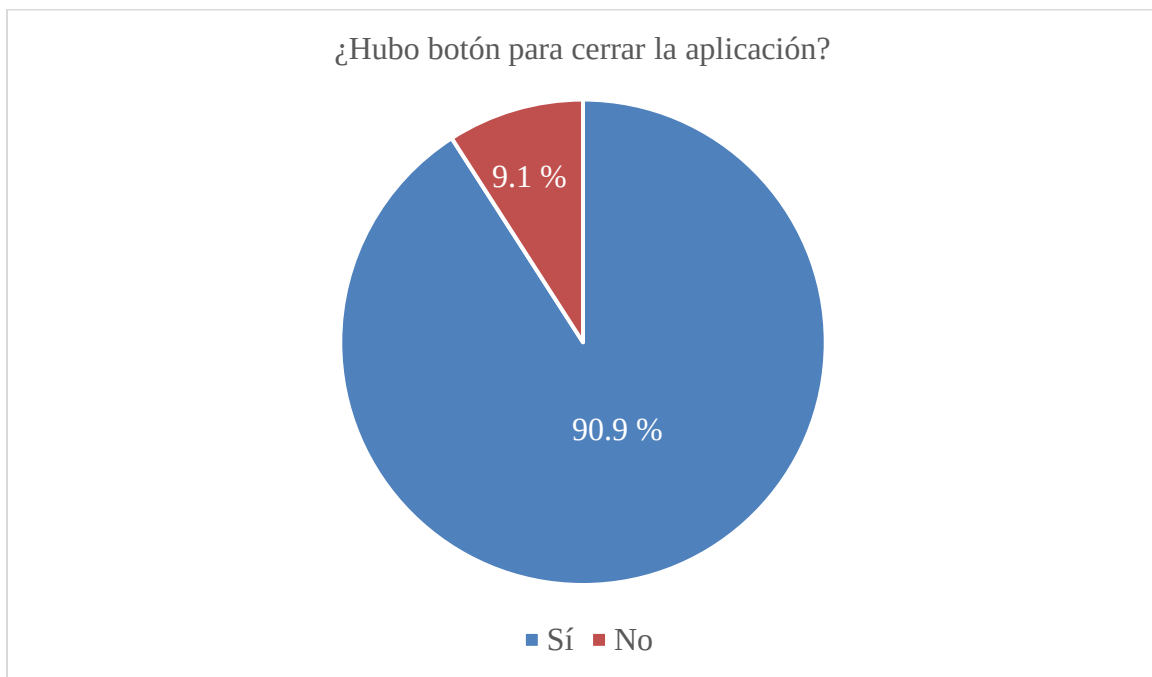
Cuando se les preguntó si la aplicación fortalece el aprendizaje del idioma inglés en la carrera de Ingeniería de Sistemas un 27.3 % considera que tal vez es importante, y el restante 72.7 % dijo que sí lo era. (Gráfica 5)



Grafica 5 ¿Considera que la aplicación fortalece el aprendizaje del idioma inglés en la carrera?

Fuente: Formularios de Google

En la parte gráfica se encuestó sobre la presencia de ciertas secciones, una de ella fue si había un botón para cerrar la aplicación, un 90.9 % dijo que, si había un botón para cerrar la aplicación, mientras un 9.1 % dijo que no había botón para cerrar la aplicación, mirar gráfica 6.



Grafica 6 ¿Hubo botón para cerrar la aplicación?

Fuente: Formularios de Google

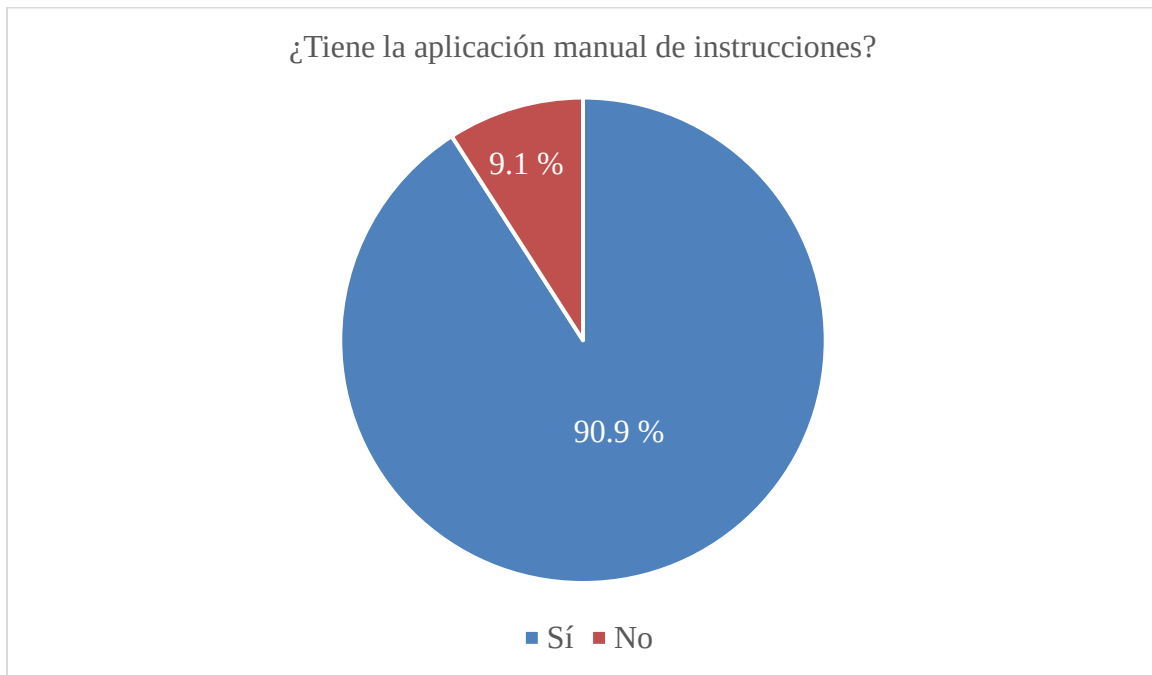
También en la parte gráfica se evaluó si era entendible la aplicación, y el 100 % dijo que si era entendible entre un 75 % y el 100 %, ver gráfica 7.



Grafica 7 ¿Fue la información entendible?

Fuente: Formularios de Google

Igualmente, se preguntó si la aplicación contaba con algún manual de instrucciones, donde, un 90.9 % dijo que sí lo había, entre tanto que un 9.1 % dijo que no había manual de instrucciones, observar gráfica 8.



Grafica 8 ¿Tiene la aplicación manual de instrucciones?

Fuente: Formularios de Google

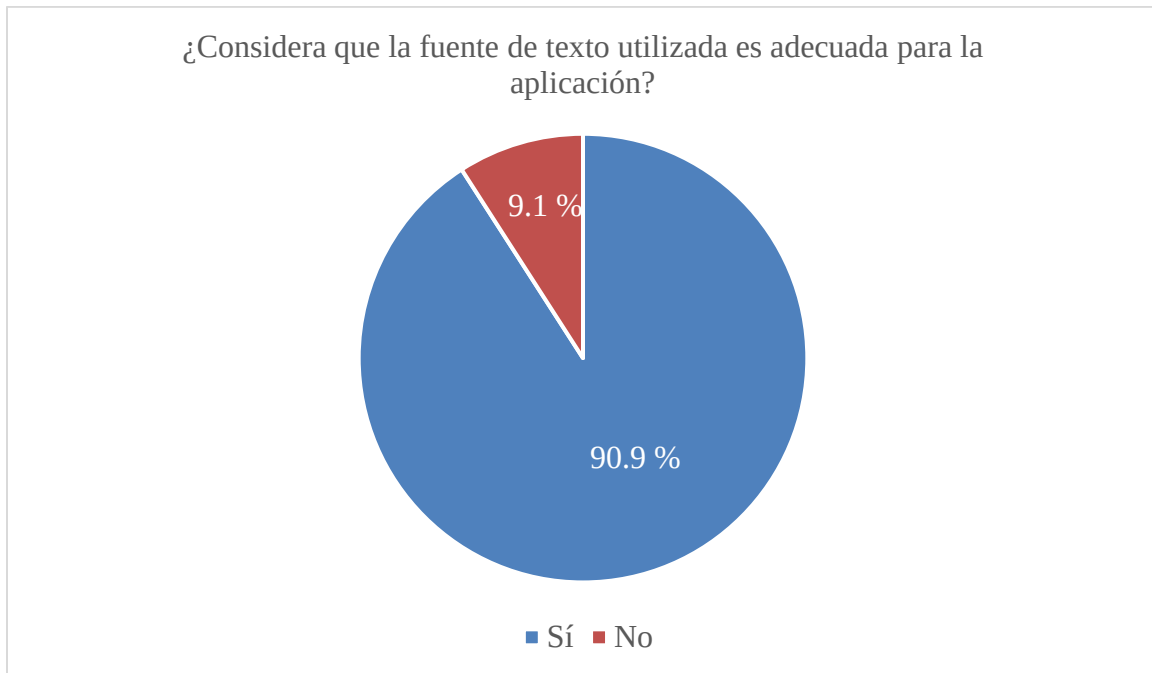
Del mismo modo, se inquirió si el texto era legible, de donde el 100 % dijo que sí, mirar gráfica 9.



Grafica 9 ¿Fue legible el texto en la aplicación?

Fuente: Formularios de Google

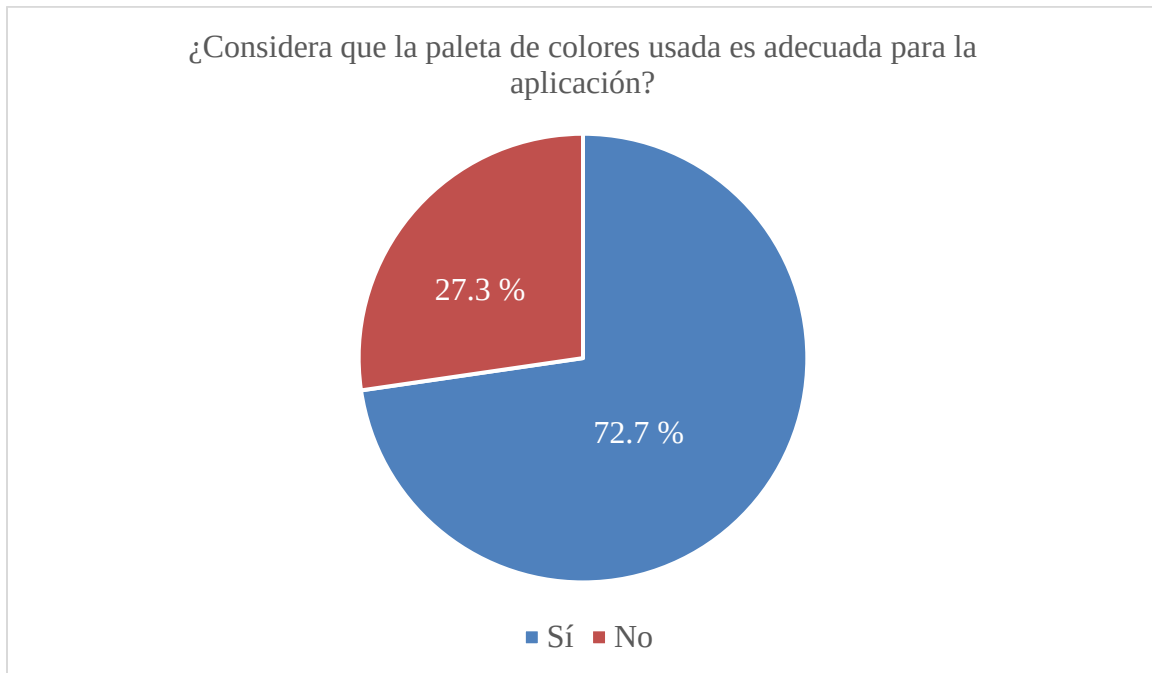
Se indagó si la fuente de texto era la adecuada para la aplicación, un 90.9 % dijo que si era adecuada, a la vez que un 9.1 % dijo que no lo era. (Gráfica 10)



Gráfica 10 ¿Considera que la fuente de texto utilizada es adecuada para la aplicación?

Fuente: Formularios de Google

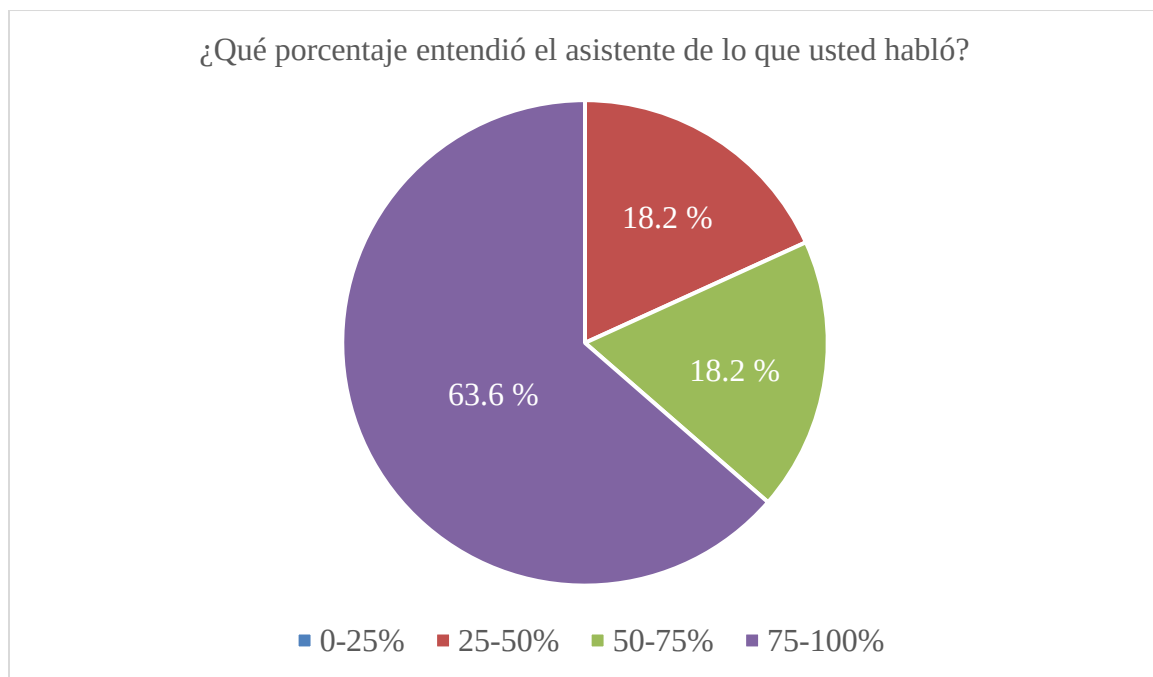
Se consultó si la paleta de colores era adecuada para la aplicación un 72.7 % respondió si, de la misma manera 27.3 % dijo que no. (Gráfica 11)



Grafica 11 ¿Considera que la paleta de colores usada es adecuada para la aplicación?

Fuente: Formularios de Google

Asimismo, se preguntó cuánto había entendido el asistente de lo hablado por los estudiantes, donde un 63.6 % dijo que había entendido entre un 50 % y un 75 %, paralelamente un 18.2 % dijo que el entendimiento fue entre un 75 % y un 100 %, añadido a esto otro 18.2 % respondió que el asistente entendió entre un 25 % y un 50 % de lo que se había dicho, ver gráfica 12.



Gráfica 12 ¿Qué porcentaje entendió el asistente de lo que usted habló?

Fuente: Formularios de Google

Añadido a lo anterior los estudiantes encuestados añadieron dentro de las recomendaciones, agregar más temas de conversación, entrenar más el asistente, corregir los niveles de dificultad, incluir más variedad de respuestas, también que el uso del micrófono indispensable para que la aplicación comprenda.

Dentro de los temas de conversación los participantes recomendaron: video juego, cine, política, economía, temas abiertos, cultura general y de Ingeniería de Sistemas.

Conclusiones

Se concluye que, de acuerdo con los datos obtenidos, que el uso de herramientas de la inteligencia artificial como asistentes virtuales, herramientas de reconocimiento del habla y sintetizadores de voz, pueden ser un apoyo fundamental al desarrollo del idioma inglés en los estudiantes de la carrera de Ingeniería de Sistemas, ya que por la versatilidad que estas herramientas ofrecen, pueden cubrir áreas de necesidad, que estén presentes en los niveles de inglés. Estos sistemas también pueden ser utilizados con otros idiomas como español, árabe, mandarín, francés, alemán, japones, coreano y portugués.

Las herramientas IBM Watson, Speech to Text y Text to Speech tienen la capacidad de aprender, por esta razón es conveniente utilizarlas, con más de un 80 % de los usuarios que manifestaron haber sido entendidos entre un 50 % y un 100 %. También cabe resaltar que el establecimiento de intenciones claras y concisas en el asistente es indispensable, ya que por esa causa el asistente tuvo algunas interrupciones con frases que eran relativamente sencillas, asimismo una buena conexión a internet es vital para el buen fluir de las comunicaciones entre el usuario y el asistente.

Estas tecnologías también se pueden utilizar en proyectos relacionados con servicio al cliente, asimismo otro tipo de usos que podría tener IBM Watson es el Internet de las cosas, pues su habilidad de encontrar intenciones le permite entender lo que un usuario diga, y enviar respuestas que cumplan con condiciones en determinado programa, el cual ejecutaría acciones en dispositivos electrónicos en algún entorno físico.

La plataforma de desarrollo Unity, fue una buena herramienta para la creación de este sistema, debido a la variedad de api's, plugins y sdk's los cuales permitieron que se pudieran integrar distintas tecnologías y servicios, prestados por otras plataformas, al proyecto.

El gestor de base de datos SQLite, al contar con plugins que permiten que sea utilizado por otras plataformas, pudo ser integrado sin inconvenientes en Unity. Este tuvo un rendimiento considerable, permitiendo que la aplicación funcionará sin interrupciones.

Recomendaciones

Se recomienda el uso de tecnologías de inteligencia artificial en los cursos de la Universidad de Pamplona, con el objetivo de satisfacer aquellas necesidades que de otra manera tendría un costo elevado, específicamente en el aprendizaje de idiomas, donde el contacto con una persona nativa de la lengua puede llegar a ser complejo, sumado a lo anterior el estado anímico del estudiante y el miedo a la vergüenza producida por equivocaciones, son un factor que reducen las posibilidades de práctica conversacional; teniendo en cuenta esto, un sistema que no es humano, no tiene cuidado de las equivocaciones que el usuario cometa, aunque posiblemente el flujo de la conversación se vea obstruido por la mala pronunciación, lo cual puede llevar a dos resultados.

El primer resultado es que el usuario se disguste y deje de usar la aplicación, en cambio el segundo sería un aumento en la capacidad de pronunciar las palabras de forma más precisa a causa de la persistencia, con el objetivo de ser entendido por el asistente, aunque teniendo en mente que estas tecnologías están aprendiendo todo el tiempo y hay una alta posibilidad de que el asistente comience a entender así la pronunciación no sea correcta, lo cual es bueno por reduce la posibilidad de deserción en el usuario. Entendiendo lo anterior se recomienda el uso de estas tecnologías y el entrenamiento de las mismas, para un buen desarrollo de las herramientas, que en el tiempo se hagan más precisas y de los estudiantes que las utilizan.

Glosario

- **Asistente virtual inteligente:** Es un agente virtual (Inteligencia artificial) que corre en un servidor o computador el cual procesa textos o diálogos humanos en búsqueda de intenciones, con el fin de entender al usuario y darle una respuesta propicia.
- **Intenciones:** Las intenciones son la clave dentro de una frase, lo que el usuario quiere. Ej: ¿Cuál es la ubicación del supermercado X? En la frase anterior la intención es el deseo de conocer la ubicación, otros ejemplos pueden ser la hora, el clima, etc.
- **Entidades:** Las entidades son objetos específicos Ej: ¿Cuál es la ubicación del supermercado X? En la frase anterior la entidad es el supermercado X, otros ejemplos pueden ser un color, una marca, una sucursal, etc.
- **IBM Watson:** Es el Asistente virtual inteligente de la corporación IBM
- **Speech to Text:** Tecnología que permite pasar de una onda de audio a un texto digital, mediante el uso de inteligencia artificial.
- **Text to Speech:** Tecnología que permite pasar de un texto digital a una onda de audio, mediante el uso de inteligencia artificial.

- Habilidades Cognitivas: Es la capacidad de procesar toda la información que llega a través de los sentidos (Pradas, 2020)
- Metodología Híbrida: Es una combinación de dos o más metodologías de desarrollo.
- SDK: Es un paquete de herramientas que permite a desarrolladores crear aplicaciones para plataformas específicas. (Lionel Valdellon, n.d.)
- Procesamiento de Lenguaje Natural (NLP): Es una rama de la inteligencia artificial que permite a las computadoras analizar información del lenguaje humano y realizar una interpretación. (SAS, n.d.)

Referencias

- Agarwal, R. (2021, September 1). ALEXA and the Technology Behind it. *Geek Culture*.
Recuperado Junio 15, 2022, de <https://medium.com/geekculture/alexa-and-the-technology-behind-it-e5c00793f85>
- Alisys, C. (2017, August 31). *De Audrey a Siri: historia y uso de los asistentes virtuales*. alisys.
Recuperado Junio 7, 2021, de <https://www.alisys.net/es/blog/de-audrey-a-siri-historia-y-uso-de-los-asistentes-de-virtuales>
- Amazon. (n.d.). Alexa [Illustration]. Recuperado Junio 7, 2021, de: <https://www.amazon.com/-/es/b?ie=UTF8&node=18354642011>
- Amberscript. (2019, July 11). *How Does Speech to Text Software Work | Amberscript*.
Recuperado Junio 15, 2022, de <https://www.amberscript.com/en/blog/how-speech-to-text-software-works/>
- Apple. (n.d.). Siri [Illustration]. Recuperado Junio 7, 2021, de: <https://www.apple.com/siri/>
- B, S. (2020, November 11). Chatbot History: The IBM Watson Assistant Chatbot. YakBots.
Recuperado Junio 30, 2021, de: <https://yakbots.com/chatbot-history-the-ibm-watson-assistant-chatbot/>
- Chansanchai, A. (2014, April 18). *Go behind-the-scenes of Cortana with Microsoft Research*.
The Official Microsoft Blog. Recuperado Junio 16, 2022, de <https://blogs.microsoft.com/blog/2014/04/18/go-behind-the-scenes-of-cortana-with-microsoft-research/>

- Corporación IBM. (2022, March 31). *IBM Watson Assistant—AI Chatbot*. Recuperado Junio 16, 2022, de <https://www.ibm.com/products/watson-assistant/artificial-intelligence>
- Cruz, C. (2019, April 24). *La evolución de los Asistentes Virtuales en 8 años*. Blogthinkbig.com. Recuperado Junio 7, 2021, de: <https://blogthinkbig.com/asistentes-virtuales-usuarios>
- Dudkin, I. (2018, December 11). *How Does Siri Work: Technology and Algorithm*. Skywell. Recuperado Junio 16, 2022, de Software. <https://skywell.software/blog/how-does-siri-work-technology-and-algorithm/>
- Education, I. C. (2021, August 17). Natural Language Processing (NLP). IBM Cloud. Retrieved May 8, 2022, de <https://www.ibm.com/cloud/learn/natural-language-processing>
- Ellis, R., Ellis, P. R., & Ellis (taalkunde), R. R. (1994). *The Study of Second Language Acquisition*. Oxford University Press.
- Fawzy, B. (2020, September 14). *The Technologies behind Google Assistant*. SCiPlanet. Recuperado Junio 15, 2022, de <https://www.bibalex.org/sciplanet/en/article/details.aspx?id=13640>
- Gold, B., Morgan, N., & Ellis, D. (2011). *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*. (Wiley Professional, Reference & Trade, Ed.; 2.^a ed.) [Libro electrónico].
- Google. (n.d.). Google Assistant [Illustration]. Recuperado Junio 7, 2021, de: <https://assistant.google.com/>

Guerrini, F. (2016, September 23). What Are The Differences Between AI, Machine Learning, NLP, And Deep Learning? Forbes. Recuperado Mayo 8, 2022, de <https://www.forbes.com/sites/quora/2016/09/23/what-are-the-differences-between-ai-machine-learning-nlp-and-deep-learning/?sh=dd0c54274fa1>

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. Recuperado Junio 17, 2022, de <https://doi.org/10.1080/02763869.2018.1404391>

How Do We Learn To Talk? (2019, April 18). Wonderopolis. Recuperado Junio 7, 2021, de: <https://wonderopolis.org/wonder/how-do-we-learn-to-talk>

IBM. (n.d.). Watson Assistant [Ilustración]. Watson Assistant. Recuperado Junio 30, 2021, de: <https://www.ibm.com/cloud/watson-assistant>

IBM. (n.d. -a). Watson Speech to Text [Ilustración]. Speech to Text IBM Watson | IBM. <https://www.ibm.com/cloud/watson-speech-to-text>

IBM Cloud Education. (2020, July 2). What is Natural Language Processing? Ibm. Recuperado Junio 30, 2021, de: <https://www.ibm.com/cloud/learn/natural-language-processing>

IBM Cloud Education. (2020a, July 2). *What is Natural Language Processing?* Ibm. Recuperado Junio 17, 2022, de: <https://www.ibm.com/cloud/learn/natural-language-processing>

IBM Cloud Education. (2020b, September 2). *What is Speech Recognition?* Ibm, Recuperado Junio 17, 2022, de: <https://www.ibm.com/cloud/learn/speech-recognition>

Introduction to Watson Assistant - IBM Developer. (2019, agosto 6). IBM Developer.

Recuperado 24 de mayo de 2022, de <https://developer.ibm.com/articles/introduction-watson-assistant/>

kanbanize. (2017, July 7). *What Is Kanban? Explained in 10 Minutes* | Kanbanize. Kanban

Recuperado Junio 1, 2021, de Software for Agile Project Management.

<https://kanbanize.com/kanban-resources/getting-started/what-is-kanban>

Kavlakoglu, E. (2020, November 12). *NLP vs. NLU vs. NLG: The differences between three*

natural language processing concepts. Watson Blog. Recuperado Junio 15, 2022, de

<https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/>

Klappenbach, A. (2022, January 7). Most Spoken Languages in the World 2022. *Busuu Blog*.

Recuperado Junio 15, 2022, de <https://blog.busuu.com/most-spoken-languages-in-the-world/>

Lambert, B. (1992, 21 marzo). *Louis Gerstman, 61, a Specialist In Speech Disorders and*

Processes. The New York Times. Recuperado 6 de junio de 2022, de

<https://www.nytimes.com/1992/03/21/nyregion/louis-gerstman-61-a-specialist-in-speech-disorders-and-processes.html>

Laura. (2021, December 2). How Does Ibm Watson Assistant Work? Programming & Design.

Recuperado Junio 6, 2022, de: <https://www.programatium.com/how-does-ibm-watson-assistant-work/#1>

Lionel Valdellon. (n.d.). *What is an SDK? Everything You Need to Know*. CleverTap. Retrieved June 6, 2022, from <https://clevertap.com/blog/what-is-an-sdk/>

Louise, C. (2020a, July 28). What Is Virtual Assistant Artificial Intelligence? Marine Management. Recuperado Mayo 8, 2022, de <https://marinemanagement.org/what-is-virtual-assistant-artificial-intelligence/>

Louise, C. (2020b, July 28). What Is Virtual Assistant Artificial Intelligence? Marine Management. Recuperado May 8, 2022, de <https://marinemanagement.org/what-is-virtual-assistant-artificial-intelligence/>

Lyden, C. (2017, October 11). What is Speech Recognition Software? How Does it Work? CallRail. Recuperado Junio 7, 2021, de <https://www.callrail.com/blog/speech-recognition-software/>

Microsoft. (n.d.). *What is Cortana?* Retrieved June 7, 2021, Recuperado Junio 7, 2021, de <https://support.microsoft.com/en-us/topic/what-is-cortana-953e648d-5668-e017-1341-7f26f7d0f825>

Microsoft. (n.d.-a). Cortana [Illustration]. Recuperado Junio 7, 2021, de <https://www.microsoft.com/en-us/cortana/>

Microsoft. (n.d.). [Logo de C#]. <https://docs.microsoft.com>

Mihaela, A. (2019, December 11). *The Impact of the Virtual Assistant (VA) on Language Classes - ProQuest*. Proquest. Recuperado Junio 7, 2021, de:

<https://www.proquest.com/openview/0b105fcc5ac3e50e8a9b056a44430f56/1?pq-origsite=gscholar&cbl=1876338>

- O'Boyle, B. (2021, November 2). *What is Siri and how does Siri work?* What Is Siri and How Does Siri Work? Recuperado Junio 17, 2022, de <https://www.pocket-lint.com/apps/news/apple/112346-what-is-siri-apple-s-personal-voice-assistant-explained>
- Oracle. (n.d.). Java [Logo]. Logo de Java. Recuperado Junio 16, 2022, de <https://www.java.com>
- Pradas, C. (2020, Junio 6). *20 HABILIDADES COGNITIVAS: Qué son, tipos, lista y ejemplos.* psicologia-online.com. Recuperado Junio 6, 2022, de <https://www.psicologia-online.com/habilidades-cognitivas-que-son-tipos-lista-y-ejemplos-4275.html>
- Python. (n.d.). Logo de Python [Logo]. Python. <https://www.python.com>
- Quintanilla, B. (2020, October 1). Idiomas más importantes para los negocios. *Academia Europea Corporativo*. Recuperado Junio 14, 2022, de <https://academiaeuropea.com/corporativo/idiomas-mas-hablados-en-el-mundo-corporativo/>
- Unity. (n.d.). *Logo de Unity* [Logo]. Unity. <https://www.unity.com>
- Rivera, D., & Machuca-Villegas, L. (2014). *Arquitectura de un asistente virtual para la producción de textos en la enseñanza y aprendizaje de idiomas*. Recuperado Junio 7, 2021, de Redalyc. <https://www.redalyc.org/pdf/2913/291333276014.pdf>
- Saksamudre, S. K., Shrishrimal, P., & Deshmukh, R. (2015). *A Review on Different Approaches for Speech Recognition System*. International Journal of Computer Applications, 115, 23-28.

SAS. (s.f.). *Qué es el Procesamiento de Lenguaje Natural—Natural Language Processing?*

Recuperado Junio 6, 2022, de https://www.sas.com/es_mx/insights/analytics/what-is-natural-language-processing-nlp.html

Sinicki, A. (2021, March 20). *What is Unity? Everything you need to know*. Android Authority.

Recuperado Junio 13, 2022, de <https://www.androidauthority.com/what-is-unity-1131558/>

Skinner, O. (2020, December 18). Text to Speech Technology: How Voice Computing is

Building a More Accessible World. Voices. Recuperado Junio 7, 2021, de <https://www.voices.com/blog/text-to-speech-technology/>

Somani, D. (2020, September 22). *How does Amazon Alexa work?* Gadgets Now. Recuperado

Junio 7, 2021, de <https://www.gadgetsnow.com/faqs/how-does-alexa-work/articleshow/78238756.cms>

Su-Hie, T., Yu-Jin, N., & Karthikeyan, J. (Eds.). (2022). *What is Artificial Intelligence?* L

Ordine Nuevo Publication

Suleimenova, Z. (2013). Speaking Anxiety in a Foreign Language Classroom in Kazakhstan.

Procedia - Social and Behavioral Sciences, 93, 1860–1868. Recuperado Junio 15, 2022, de <https://doi.org/10.1016/j.sbspro.2013.10.131>

Tillman, M. (2021, April 8). *What is Google Assistant and what can it do?* Pocket-Lint.

Recuperado Junio 7, 2021, de <https://www.pocket-lint.com/apps/news/google/137722-what-is-google-assistant-how-does-it-work-and-which-devices-offer-it>

Trivedi, A., Pant, N., Shah, P., Sonik, S., & Agrawal, S. (n.d.). *Speech to text and text to speech recognition systems-Areview*. 9.

Virtual Agent - IBM Watson Assistant | IBM. (2018, marzo 17). IBM. Recuperado 24 de mayo de 2022, de <https://www.ibm.com/products/watson-assistant>

Voximplant. (2020, October 1). *Blog | Voximplant.com | Voximplant.com*. What Is Automatic Speech Recognition? Recuperado Junio 14, 2022, de <https://voximplant.com/blog/what-is-automatic-speech-recognition>

Woodford, C. (2021, May 24). *How speech synthesis works*. Explain That Stuff. Recuperado Junio 15, 2022, de <http://www.explainthatstuff.com/how-speech-synthesis-works.html>

ANEXOS

**EVIDENCIAS DEL USO DEL ASISTENTE VIRTUAL INTELIGENTE CON
ESTUDIANTES INGLES DEL NIVEL I Y III DEL PROGRAMA DE SISTEMAS DE LA
UNIVERSIDAD DE PAMPLONA AMPLIACIÓN VILLA DEL ROSARIO**

YONATHAN JAVIER JAIMES MANTILLA



UNIVERSIDAD DE PAMPLONA

FACULTAD DE INGENIERÍAS Y ARQUITECTURA

**DEPARTAMENTO DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA, SISTEMAS Y
TELECOMUNICACIONES**

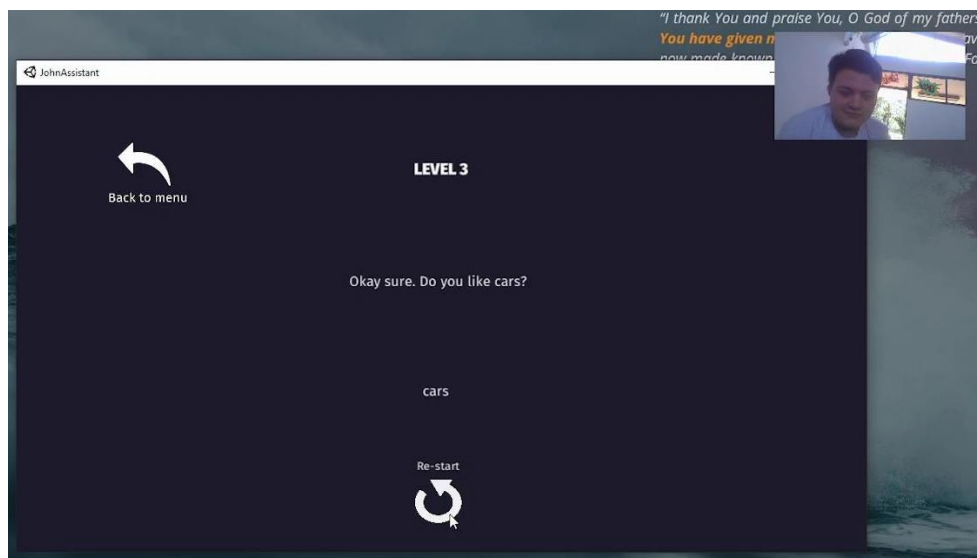
INGENIERÍA DE SISTEMAS

VILLA DEL ROSARIO

2021



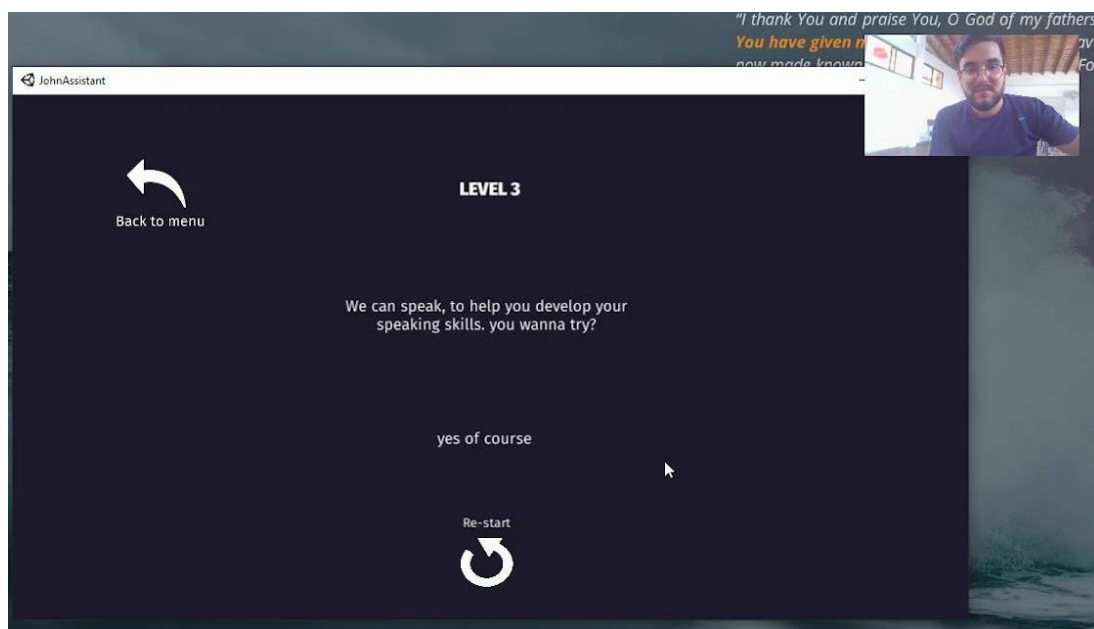
Evidencia 1 Estudiante de Nivel I de Ingles del grupo BR probando la aplicación



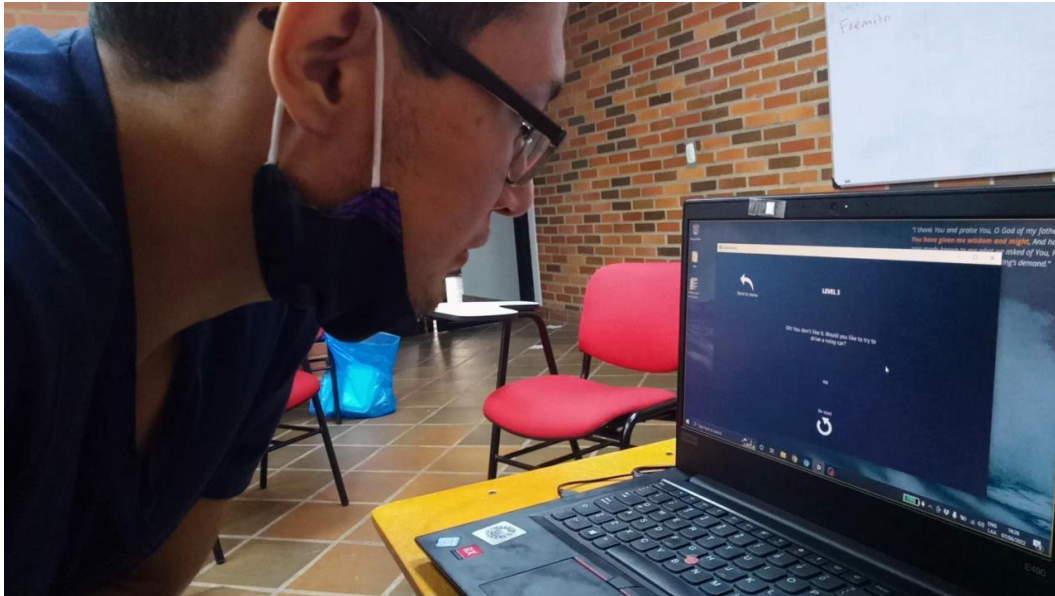
Evidencia 2 Estudiante de Nivel I de Ingles del grupo BR probando la aplicación



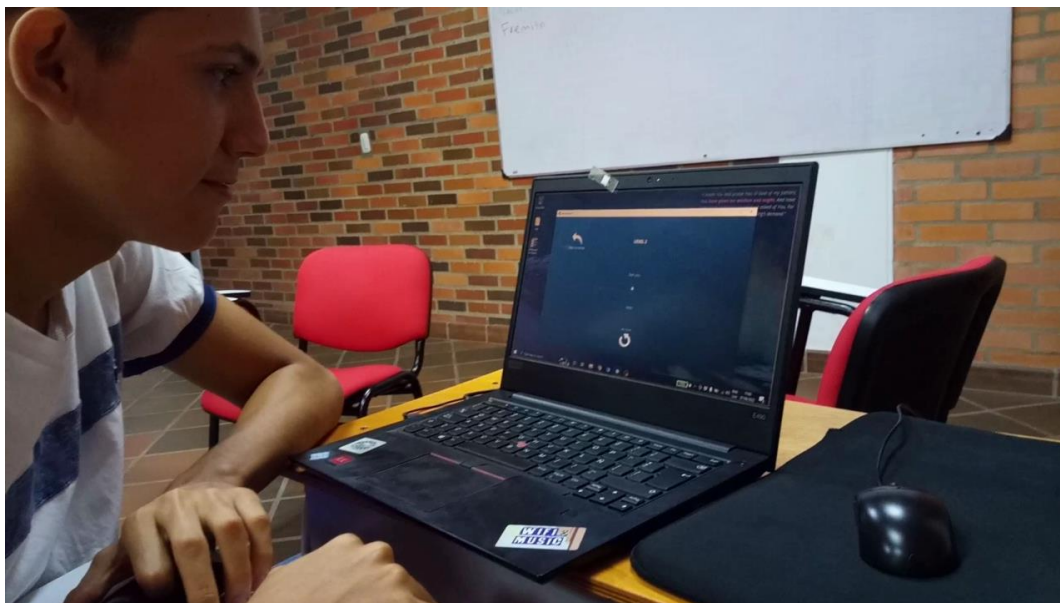
Evidencia 3 Estudiante de Nivel I de Ingles del grupo BR probando la aplicación



Evidencia 4 Estudiante de Nivel I de Ingles del grupo BR probando la aplicación



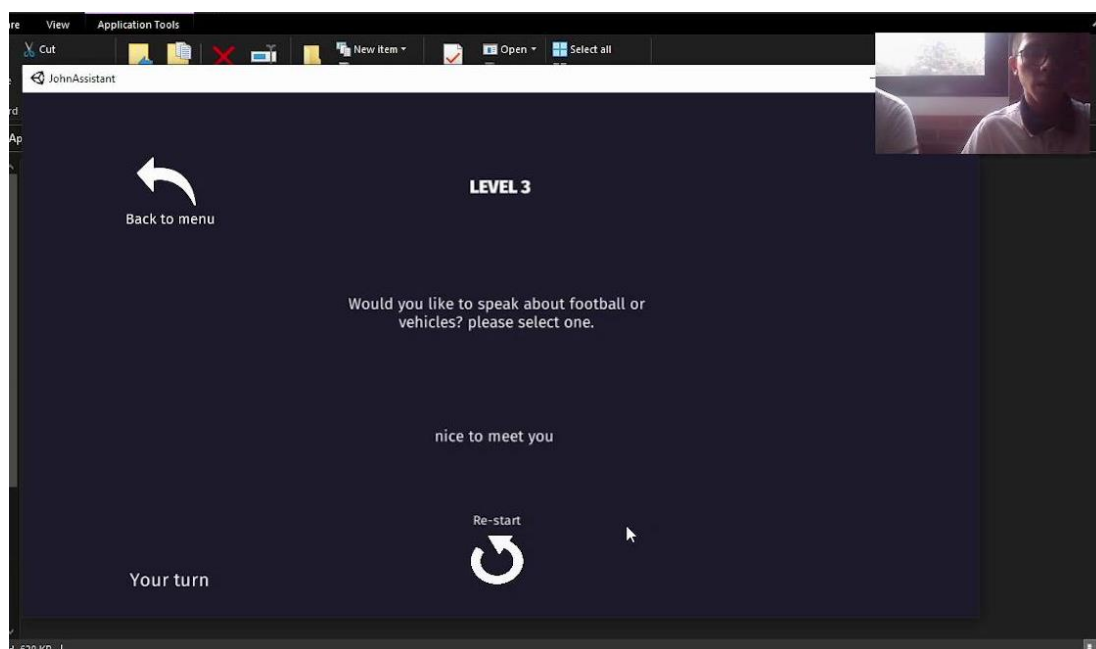
Evidencia 5 Estudiante de Nivel I de Ingles del grupo AR probando la aplicación



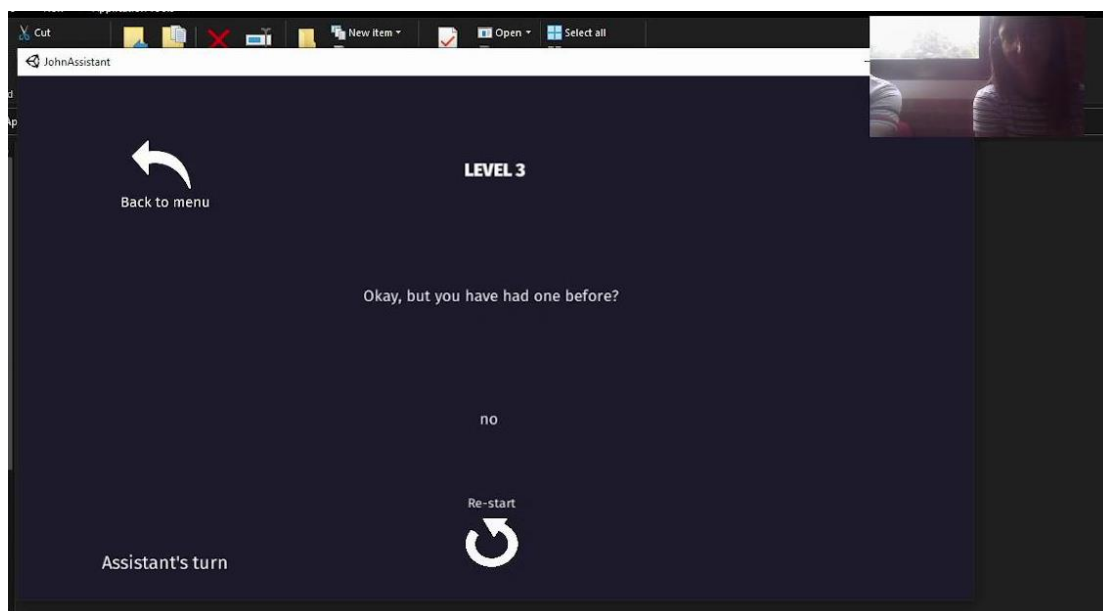
Evidencia 6 Estudiante de Nivel I de Ingles del grupo AR probando la aplicación



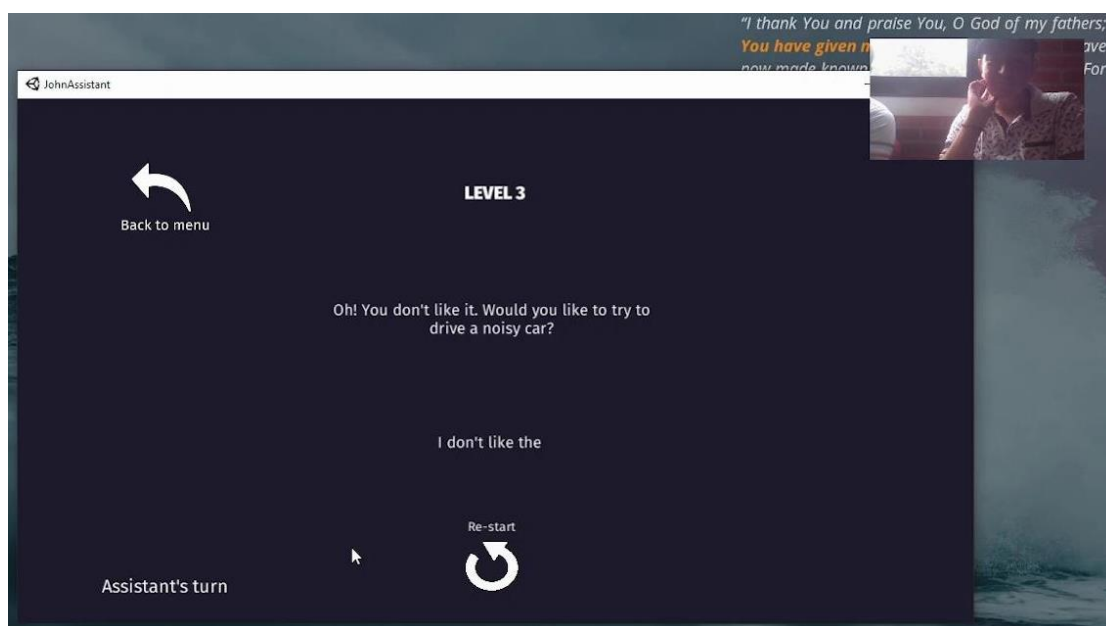
Evidencia 7 Estudiante de Nivel I de Ingles del grupo AR probando la aplicación



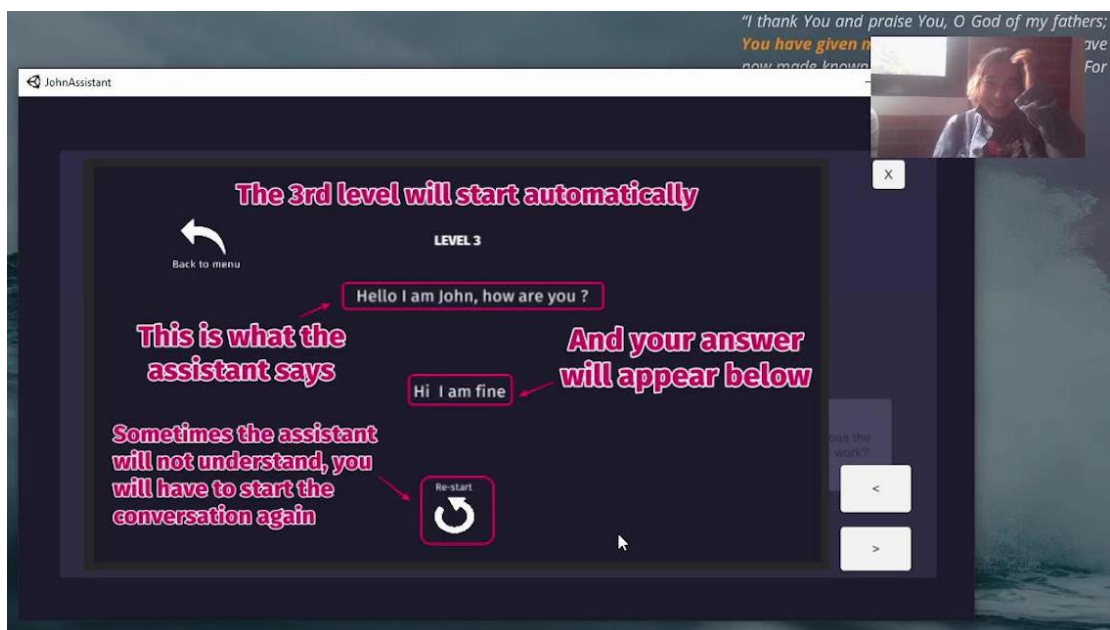
Evidencia 8 Estudiante de Nivel III de Ingles del grupo AR probando la aplicación



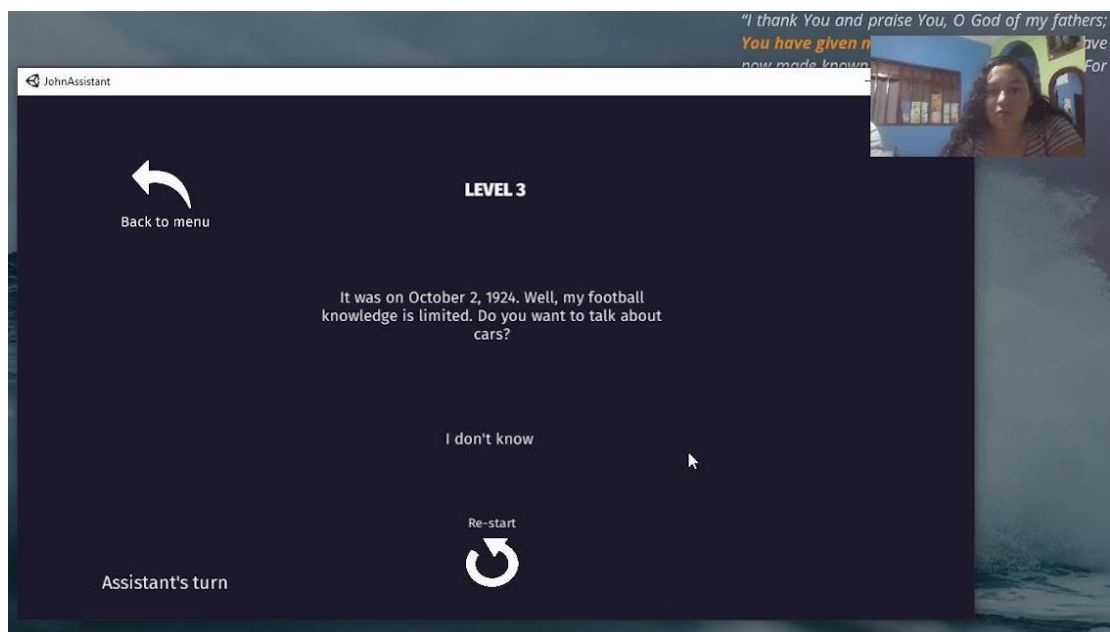
Evidencia 9 Estudiante de Nivel III de Ingles del grupo AR probando la aplicación



Evidencia 10 Estudiante de Nivel III de Ingles del grupo AR probando la aplicación



Evidencia 11 Estudiante de Nivel III de Ingles del grupo AR mirando el manual de la aplicación



Evidencia 12 Estudiante de Nivel III de Ingles del grupo AR probando la aplicación