



## Pronuncia español con Sarahí



Informe técnico

Autor: Tayde Martín Cruz Lovera

31 de enero de 2020

## Contenido:

Informe técnico	1
Contenido:	1
DESCRIPCIÓN	2
ANÁLISIS	2
Definición del proyecto	2
Requerimientos de la aplicación	3
Alcance	3
PLANEACIÓN	4
Recursos humanos y roles	4
Recursos técnicos	4
Herramienta para la programación de la aplicación	5
Cronograma de actividades	6
DISEÑO	7
Diseño de la interacción	7
Diseño conceptual de la interfaz de usuario	8
Búsqueda de soluciones	10
Descripción de los niveles de la aplicación	11
DESARROLLO	11
Mejoras en el modelo del avatar virtual (Sarahi)	11
Programación del prototipo	12
Diseño de la interfaz de usuario	13
Programación de la práctica 1	14
Programación de la práctica 2	15
Programación de la práctica 3	16
Programación de la práctica 4	17
Programación de la sección de órganos del AFH	18
Programación del juego 1	19
Programación del juego 2	20
Programación del examen	21
Corrección de errores	23
PRUEBAS CERRADAS CON UN GRUPO DE USUARIOS	24
LANZAMIENTO DE LA APLICACIÓN EN LAS TIENDAS	24
Requisitos para la instalación de la aplicación	24

## DESCRIPCIÓN

Pronuncia español con Sarahí es una aplicación interactiva que permite practicar sonidos del español. Está conformada por la descripción de los órganos del aparato fonador, cuatro prácticas de algunos sonidos problemáticos en el aprendizaje del español: vocales, diptongos, distinción entre /b/ y /p/, /d/ y /t/ y /r/; trabalenguas, dos juegos y un examen.

Es posible visualizar el movimiento de los órganos que intervienen en el habla con el objetivo de ayudar a practicar y mejorar la pronunciación en español.

La aplicación Pronuncia español con Sarahí es parte del proyecto PAPIME PE404019. El aparato fonador humano en tercera dimensión: Sarahí 2, de la Universidad Nacional Autónoma de México.

Esta aplicación te ayuda a practicar las vocales, los diptongos, las diferentes realizaciones del sonido /r/: simple y vibrante, reconocer la diferencia entre los sonidos /b/ y /p/ y reconocer la diferencia entre los sonidos /d/ y /t/.

La aplicación tiene las siguientes características:

- Es una herramienta que permite practicar la pronunciación del español en cualquier momento y en cualquier lugar de una manera educativa y entretenida.
- Se puede visualizar los movimientos que se llevan a cabo al pronunciar los sonidos mencionados anteriormente.
- Consta de cuatro prácticas, dos juegos y un examen global para comprobar los avances.
- Permite escuchar, visualizar el aparato fonador humano y grabar voz.

## ANÁLISIS

### Definición del proyecto

La aplicación “Pronuncia español con Sarahí” tiene como antecedente el proyecto “El aparato fonador humano en tercera dimensión” desarrollado en el 2005, la cual se diseñó para su uso en el Observatorio Ixtli y su infraestructura de realidad virtual. Como punto de partida para el desarrollo de esta aplicación se necesitó recuperar los recursos generados en el proyecto “El aparato fonador humano en tercera dimensión” (2005), es decir, modelos 3d, animaciones, texturas y audios.

La aplicación “Pronuncia español con Sarahí” permite al usuario realizar cuatro prácticas, dos juegos y un examen de escritura y pronunciación de la lengua española, además se muestra el modelo 3D de una mujer, con animaciones y audios de sílabas, palabras, oraciones y trabalenguas dependiendo del ejercicio. Para cada práctica, la aplicación tiene un número determinado de ejercicios con sus respectivos incisos, que deberán ser resueltos por el usuario.

Está diseñada para funcionar en modo vertical (portrait) en dispositivos móviles, celulares y tabletas, con sistema operativo Android 4.1 o superior o bien iOS 9.0 o superior. Toda interacción con la aplicación se realiza mediante gestos táctiles en la pantalla del dispositivo móvil.

Para cada ejercicio y juego, se puede rotar el modelo 3D, alejar y acercar, así como realizar un movimiento vertical y horizontal, también es posible ver el interior del modelo 3D o bien centrar la cámara en la lengua y paladar del modelo 3D. Para poder acceder al examen, el alumno deberá haber realizado correctamente un 30% de los ejercicios de cada práctica.

Adicionalmente, la aplicación cuenta con una sección donde se muestran las partes del aparato fonador humano, así como la información de cada órgano en forma de texto y audio. La aplicación puede mantener un registro en el dispositivo del avance de un solo usuario, sin embargo, la información se elimina al ejecutar nuevamente la aplicación.

#### Requerimientos de la aplicación

- Plataforma destino: dispositivos móviles con distribución en tiendas oficiales de aplicaciones (Android y iOS).
- Usuario objetivo: alumnos extranjeros, aunque también la pueden aprovechar niños con problema del habla.
- Debe poder ser utilizada por el alumno solo o con el profesor.
- Se contemplan únicamente las 4 prácticas de la aplicación original.
- En las prácticas se debe mostrar el avatar virtual (Sarahí) para que el usuario observe las animaciones de la pronunciación al tiempo que escucha el audio correspondiente.
- Debe incluir una práctica final (examen), accesible hasta que el usuario haya completado correctamente un 30% de los ejercicios de cada una de las 4 prácticas.
- Debe incluir dos juegos simples.
- Debe incluir al igual que en la aplicación original, una parte en donde se muestren los órganos del aparato fonador humano con descripción en texto y audio.
- En las prácticas debe incluir la opción de ver los órganos en el interior del avatar virtual (Sarahí).
- En las prácticas s debe permitir la manipulación de la rotación, posición y acercamiento del avatar virtual.
- Cada práctica debe tener una pantalla de presentación.
- Sonido ambiental en el menú principal.
- Es posible avanzar entre ejercicios sin haberlos respondido.
- Debe incluir una venta de créditos y una ventana de ayuda con instrucciones de uso.

#### Alcance

Para el desarrollo de la aplicación se reutilizaron los recursos generados para el proyecto “El aparato fonador humano en tercera dimensión” (2005), es decir, modelos 3D, texturas, audios y animaciones. Los modelos 3D y las texturas se mejoraron para mejorar su calidad, pero respecto

a las animaciones y audios no se generaron más de forma adicional a los que fue posible recuperar de los respaldos del proyecto “El aparato fonador humano en tercera dimensión” (2005). La aplicación se desarrolló para ser ejecutada en dispositivos móviles (Android y iOS) únicamente, la forma de distribución a los usuarios es mediante las tiendas oficiales de Android (Google Play) y iOS (App Store) de forma gratuita.

## PLANEACIÓN

### Recursos humanos y roles

1. Dra. Rosa Esther Delgadillo Macías – Coordinación y responsable del proyecto (CEPE)
2. Mtra. Estefany Isabel Macías Rioja – Corresponsable (CEPE)
3. Mat. María del Carmen Ramos Nava – Administración del proyecto (DGTIC)
4. Tayde Martín Cruz Lovera – Coordinador técnico (DGTIC)
5. Ing. José Martín Arreola Manzo – Programación de interacción (DGTIC)
6. Arq. Ricardo Cornejo Bravo – Modelado y animación (DGTIC)
7. José Adrián Romero Jiménez – Diseño de interfaz gráfica (DGTIC)
8. Ing. José Larios Delgado – Arquitecto de soluciones 3D (DGTIC)
9. D.C.V. Víctor Hugo Franco Serrano – Arquitecto de soluciones 3D (DGTIC)

### Recursos técnicos

1. Computadora para la programación y compilación de la aplicación en Android con las siguientes características:
  - Unity Hub 2.2.0 o posterior y Unity versión: 2019.2.0f1 con los módulos de Android y iOS
  - Android SDK, NDK y JDK
  - Windows 7 o superior
  - Tarjeta de video dedicada NVIDIA GT 730 o superior, AMD Radeon R7 240 o superior
  - Mínimo 8 GB de memoria RAM
2. Computadora para compilación de la aplicación en iOS con las siguientes características:
  - Xcode 11.2 o superior
  - Computadora Apple
  - MacOS Mojave 10.4.6 o Superior
3. Computadora para modelado 3D y animación con las siguientes características:
  - Autodesk Maya
  - Autodesk Mudbox
  - Adobe Photoshop
  - Windows 7 o superior
  - Tarjeta de video dedicada NVIDIA GTX 1050 Ti o superior
  - Mínimo 16 GB memoria RAM o superior
4. Computadora para diseño de la interfaz con las siguientes características:
  - Adobe Illustrator

- 
- Windows 7 o superior
  - Tarjeta de video dedicada NVIDIA GT 730 o superior
  - Mínimo 8 GB de memoria RAM (se recomiendan 16 GB)

Herramienta para la programación de la aplicación

En el Departamento de Visualización y Realidad Virtual se analiza para cada proyecto la mejor opción para el desarrollo dependiendo de la plataforma destino y los requerimientos. En general las principales opciones actualmente son:

- Unity
- Unreal
- OpenSceneGraph (OSG)

Para el desarrollo de esta aplicación en específico se optó por utilizar Unity por ser un proyecto destinado a dispositivos móviles en donde Unity tiene una gran fluidez de ejecución, además de permitir empezar a desarrollar de forma ágil y rápida.

Unity – Herramientas para desarrollo de aplicaciones gráficas interactivas con soporte de compilación para distintas plataformas como PC, web, Android, iOS, VR/AR, etc.

Ventajas de su uso sobre otras herramientas similares:

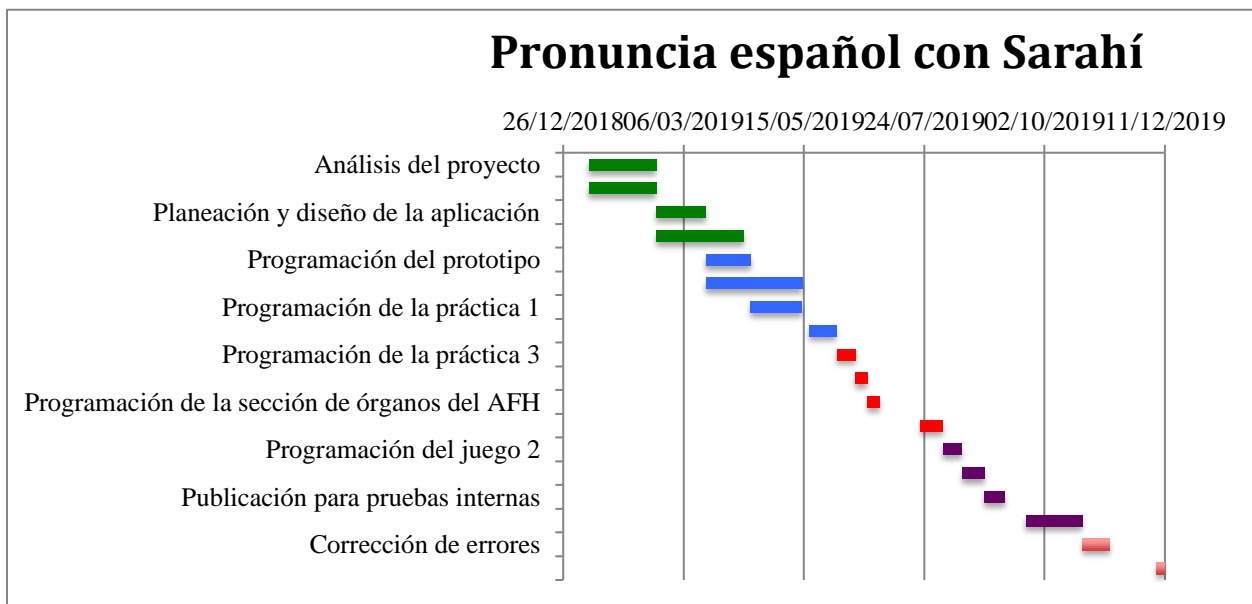
- Curva de aprendizaje fácil.
- Facilidad y rapidez para el desarrollo.
- Cantidad de documentación y tutoriales.
- Comunidad grande de usuarios y muy activa.
- Gran variedad de contenido de terceros que facilitan la integración de dispositivos o características.
- Soporte de compilación para una gran variedad de plataformas.

Desventajas:

- Requiere licencia.
- No es de código abierto.
- Presencia de errores que se van solucionando con parches frecuentes.

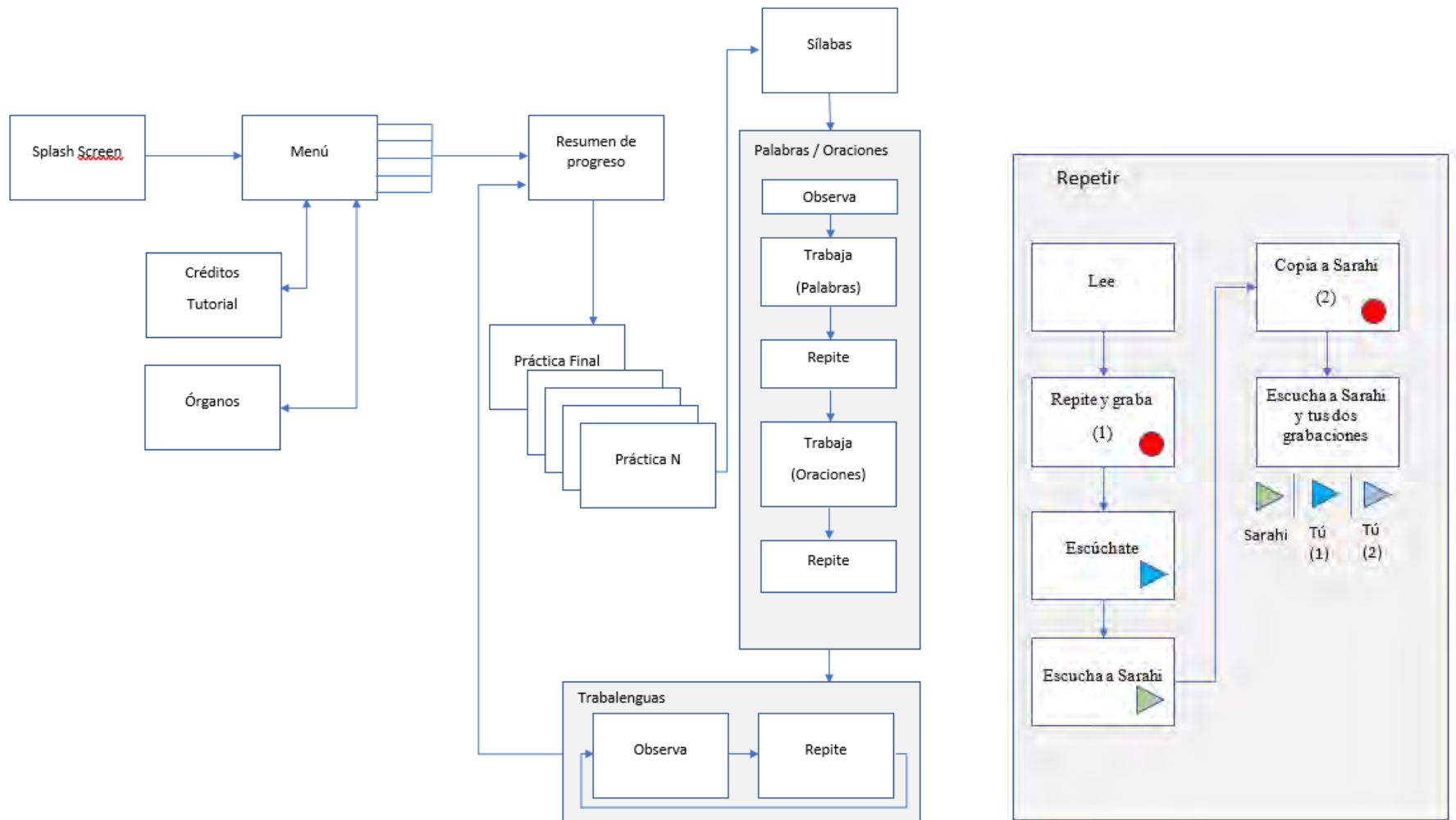
Cronograma de actividades

Tarea	Fecha de inicio	Fecha final	Duración (días)
Análisis del proyecto	10/01/2019	18/02/2019	39
Búsqueda de soluciones	10/01/2019	18/02/2019	39
Planeación y diseño de la aplicación	18/02/2019	19/03/2019	29
Mejoras en el modelo del avatar virtual (Sarahí)	18/02/2019	10/04/2019	51
Programación del prototipo	19/03/2019	14/04/2019	26
Diseño de la interfaz de usuario	19/03/2019	14/05/2019	56
Programación de la práctica 1	14/04/2019	14/05/2019	30
Programación de la práctica 2	18/05/2019	03/06/2019	16
Programación de la práctica 3	03/06/2019	14/06/2019	11
Programación de la práctica 4	14/06/2019	21/06/2019	7
Programación de la sección de órganos del AFH	21/06/2019	28/06/2019	7
Programación del juego 1	22/07/2019	04/08/2019	13
Programación del juego 2	04/08/2019	15/08/2019	11
Programación del examen	15/08/2019	28/08/2019	13
Publicación para pruebas internas	28/08/2019	09/09/2019	12
Inicio de pruebas con usuarios	21/09/2019	24/10/2019	33
Corrección de errores	24/10/2019	09/11/2019	16
Lanzamiento de la aplicación en las tiendas	06/12/2019	11/12/2019	5



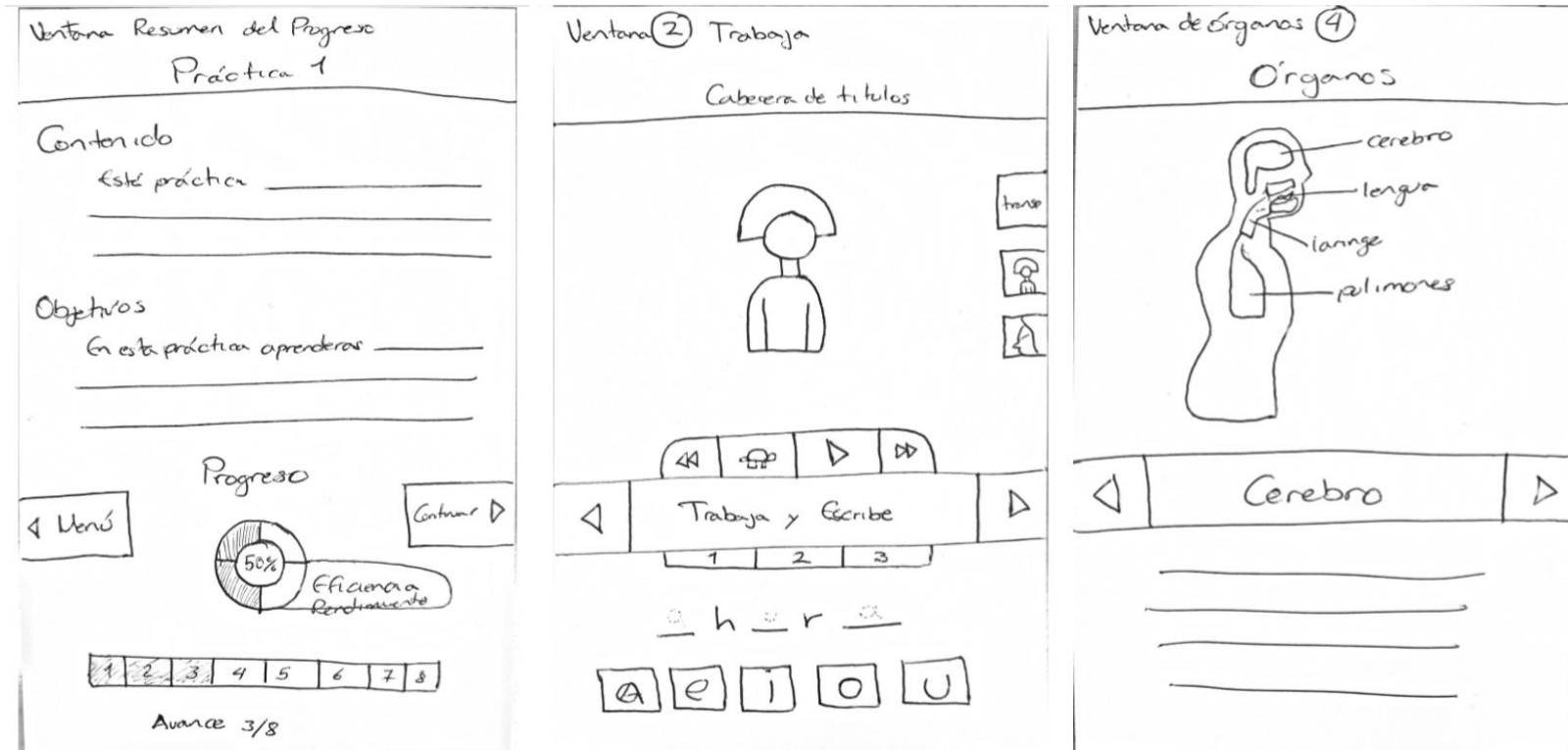
## DISEÑO

### Diseño de la interacción





Diseño conceptual de la interfaz de usuario



Cabecera de Titulos	PALABRAS	ORGANOS
<p>Despliegue 3D</p> <p>Controles</p>	 <p>Controles de reproducción</p>	
<p>Zona de Instrucciones</p>	<p>Completa la palabra</p>	<p>Cuerdas Vocales</p>
<p>Zona de Interacción</p>	<p>_ H _ R _</p> <p>A C F D H M O</p> <p>U S G T O P</p>	<p>Las cuerdas vocales humanas son la parte del aparato fonador directamente responsable de la producción de la voz. No tienen forma de cuerda, sino que se trata de una serie de repliegues o labios membranosos por lo que es un error llamarlas cuerdas.</p>

## Búsqueda de soluciones

### - Animaciones

Para la reutilización de los modelos y animaciones recuperadas se presentaron distintas opciones:

- Utilización directa de las animaciones y modelos obtenidos del ejecutable de la aplicación en formato de CAL3D para utilizarlas en Unity, esto requería programar el código que convirtiera del formato de CAL3D a modelos y animaciones de Unity.
- Conversión de las animaciones y modelos obtenidos del ejecutable de la aplicación en formato de CAL3D a otro formato como FBX para posteriormente importarlas dentro de Unity.
- Exportación de las animaciones y modelos de los archivos fuente a formato FBX para posteriormente importarlas dentro de Unity.

Después de analizar las opciones se optó por utilizar los archivos fuente para tener la libertad de realizar cambios en el modelo y texturas para mejorar su calidad, cosa que en las otras opciones no era posible.

### - Reconocimiento de voz

Se propuso utilizar alguna herramienta para reconocimiento de voz para evaluar la pronunciación de las palabras, frases y trabalenguas en la aplicación por parte del usuario, esto llevó a la necesidad de investigar las posibles opciones compatibles con Unity para integrarlas en el desarrollo:

- Google
- Microsoft
- IBM
- Amazon
- Otras

La dificultad de utilizar estas herramientas es que la mayoría tienen un costo dependiendo del uso del sistema de reconocimiento de voz que llevaría a la necesidad de tener recursos económicos para mantener el servicio, otras alternativas menos populares tanto de costo como gratuitas requerían más tiempo de pruebas para garantizar su integración con Unity, su correcto funcionamiento con el idioma español y su viabilidad como herramienta de evaluación de pronunciación. Debido a esto se decidió no integrarlo y simplificar esa función, quedando como una opción de autoevaluación por parte del usuario grabándose, escuchándose y comparándose entre el audio grabado por el usuario y el audio de la pronunciación del avatar virtual de la aplicación (Sarahi).

---

## Descripción de los niveles de la aplicación

La aplicación consta de los siguientes niveles:

- Splash DVRV – Muestra una presentación de los logos de las partes involucradas en el desarrollo del proyecto. Este nivel solo se ejecuta una vez al iniciar la aplicación.
- Menú – Esta escena contiene los botones para acceder a cada una de las prácticas de la aplicación, así como a la sección de órganos y a los juegos.
- Prácticas – Esta es la escena principal donde están implementadas las 4 prácticas de la aplicación, así como el examen. Dependiendo de los parámetros iniciales al iniciar la escena, se muestran los datos correspondientes a la práctica seleccionada desde el menú.
- Órganos – En esta escena se encuentra el modelo 3D de los órganos del aparato fonador humano. Se muestran los nombres de las partes y los botones de cada parte permiten mostrar la información correspondiente de cada una.
- Juego1 – En esta escena está implementado el juego de relación de columnas. Se relacionan imágenes con palabras escritas, usando botones en la UI.
- Juego2 – En esta escena está implementado el juego de relación de columnas. Se relacionan enunciados con las palabras que las completan correctamente, usando botones en la UI.

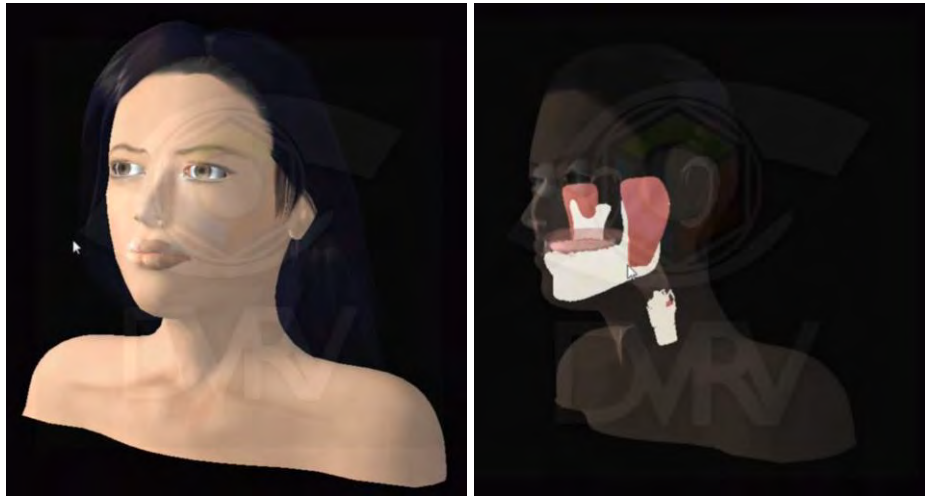
## DESARROLLO

Mejoras en el modelo del avatar virtual (Sarahi)

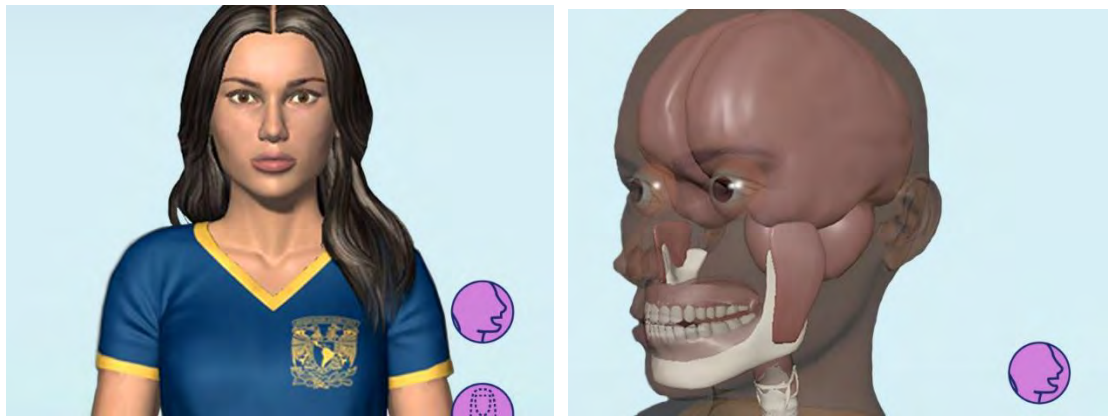
En los archivos originales recuperados se realizaron cambios en el avatar virtual con el fin de mejorar su calidad gráfica.

El modelado se llevó a cabo en varias partes, empezando con la geometría base utilizando el software Maya. Para crear a Sarahi, se comenzó a partir de algunas referencias reales para poder tener las proporciones correctas y que el personaje fuera creíble; el primer paso fue hacer una geometría con bajo número de polígonos teniendo las proporciones y rasgos básicos sin entrar a demasiado detalle, esto para facilitar los ajustes que se iban realizando sobre la marcha, también fue importante mantener la geometría de la boca lo más similar al modelo original para que las animaciones y el sistema de huesos fueran lo más compatibles posible. Después, se llevó la geometría base al software Mudbox, donde se le agregaron detalles finos como arrugas, poros de la piel, etc. utilizando un modelo de alto conteo de polígonos, ahí mismo también se generaron texturas de color y luz; con este paso terminado se extrajeron las texturas y se aplicaron al modelo original. Para finalizar, se hizo un skinning del modelo en donde se asignaron los huesos al modelo para que estos pudieran controlar la geometría y así utilizar las animaciones previamente rescatadas del proyecto original.

Mallas y texturas originales:



Mallas y texturas mejoradas:



Programación del prototipo

Se desarrolló una versión prototipo en el motor de juegos Unity para mostrar la funcionalidad principal de la aplicación a las responsables del proyecto, el cual se tomó como base para seguir el desarrollo realizando las modificaciones necesarias para mejorar la usabilidad y la experiencia de usuario. Las funciones que se programaron son las siguientes:

- Configuración de las prácticas en base a un archivo XML con la lista de palabras, frases y trabalenguas.
- Menú para acceder a las prácticas.
- Reproducción de animaciones y audios sincronizados.
- Interacción con el avatar virtual, es decir, traslación, rotación y transparencia.
- Ejercicio de completar palabras.

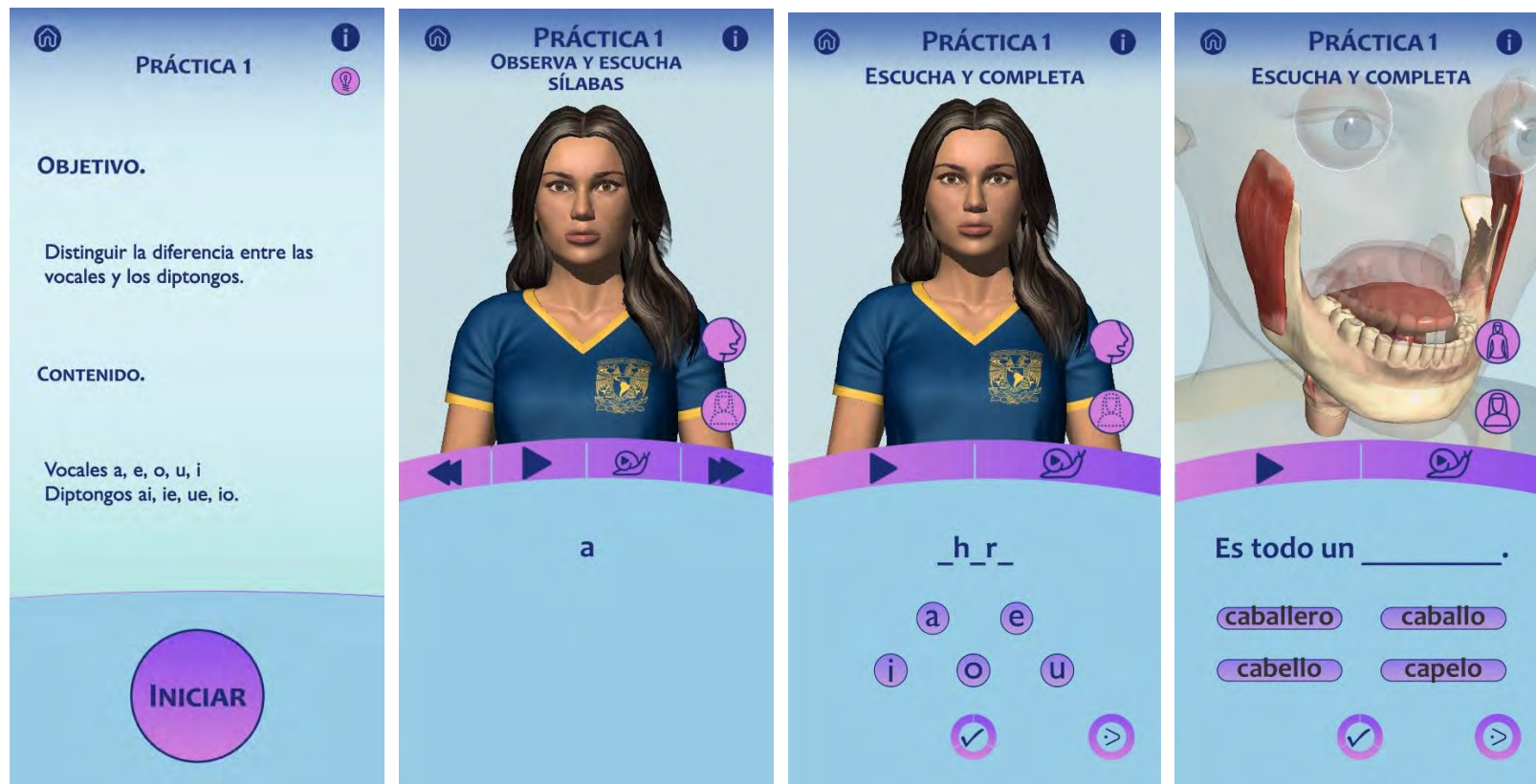
Diseño de la interfaz de usuario

A partir del diseño conceptual de la interfaz realizados se diseño una interfaz de usuario. En primer lugar, se realizó un diseño inicial que, a lo largo de gran parte del tiempo de desarrollo, se fue modificando para mejorar la usabilidad y experiencia de usuario.



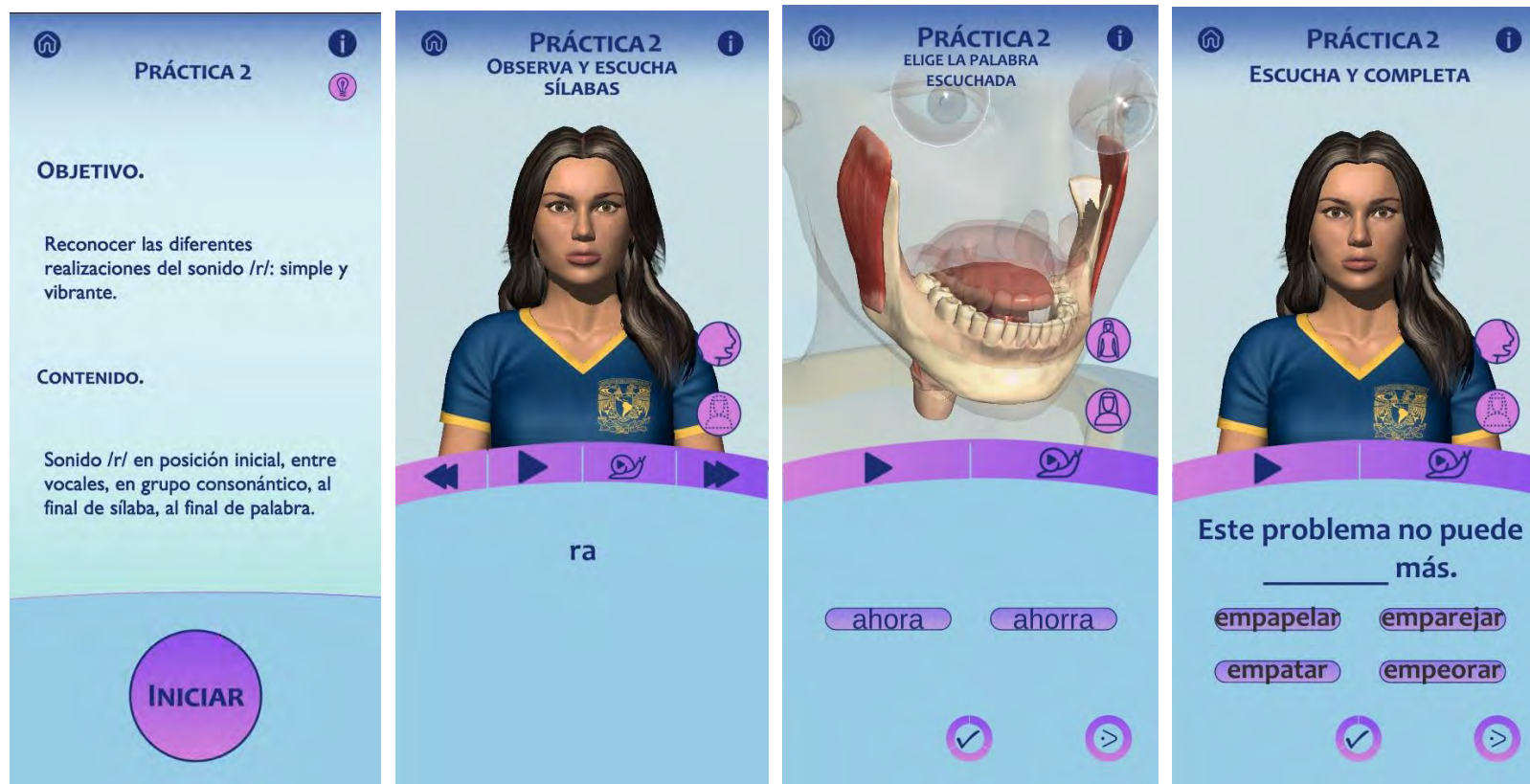
### Programación de la práctica 1

La programación de la práctica 1 se realizó tomando como base el prototipo, se implementaron los distintos ejercicios que forman parte de la práctica 1. Adicionalmente en esta fase se configuro el total de palabras, frases y trabalenguas de esta práctica, lo cual incluye audios, animaciones y textos.



Programación de la práctica 2

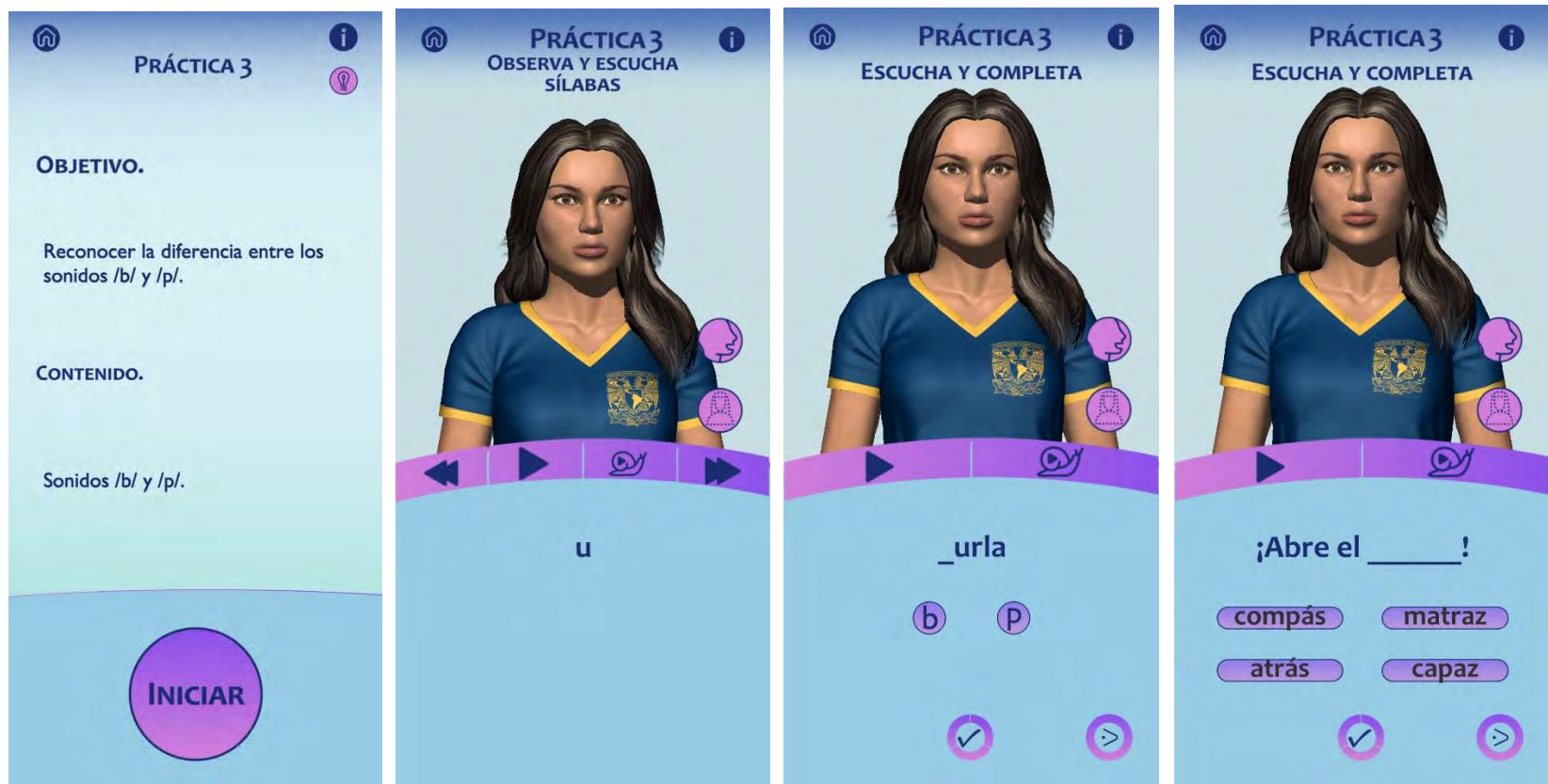
Nuevamente, la similitud entre prácticas permitió tomar como base para la programación de la práctica 2 lo realizado para la práctica 1 realizando las modificaciones específicas para esta práctica. Adicionalmente en esta fase se configuro el total de palabras, frases y trabalenguas de esta práctica, lo cual incluye audios, animaciones y textos.





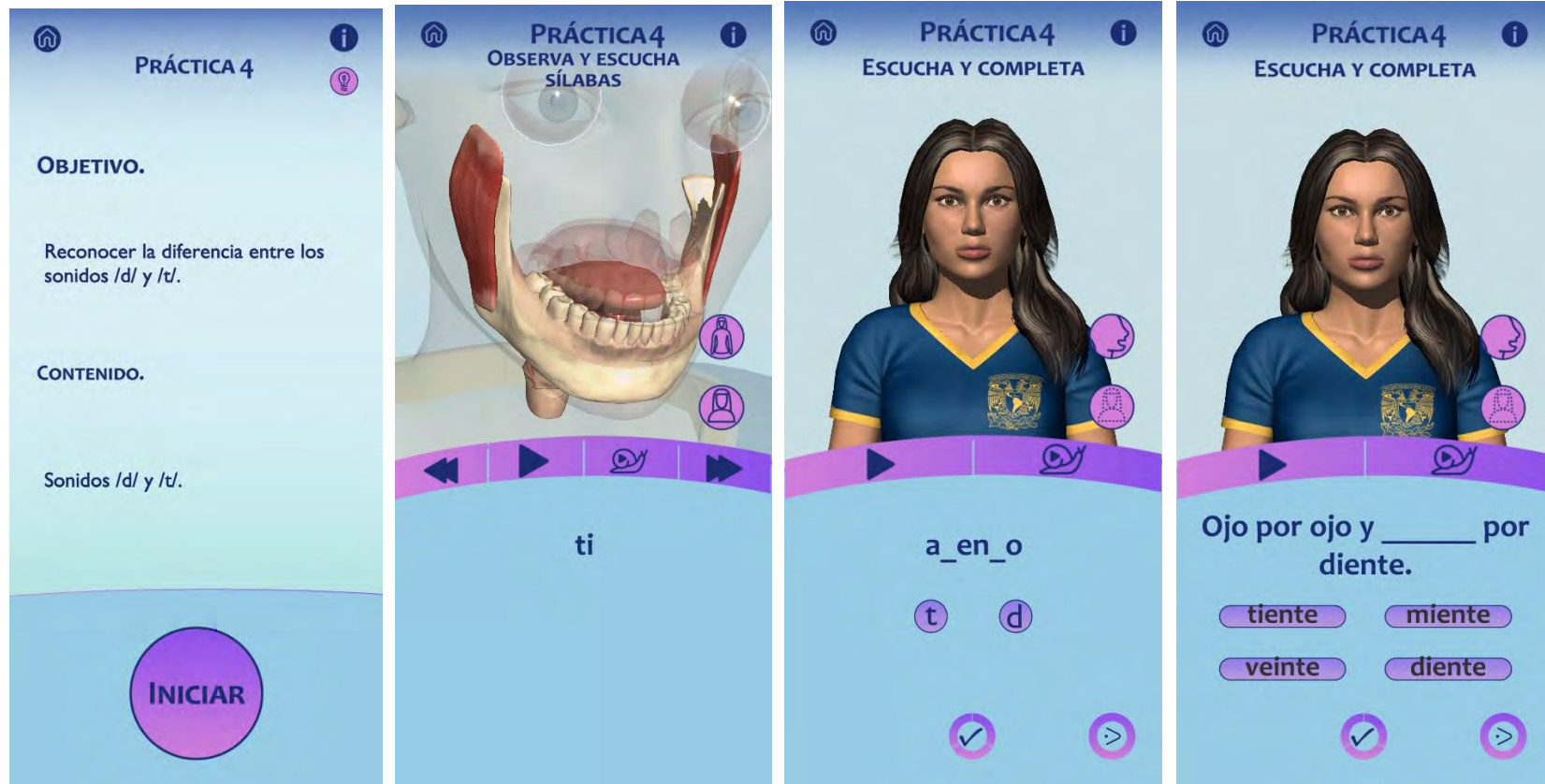
### Programación de la práctica 3

Caso similar a la fase anterior, se tomó como base para la programación de la práctica 3 lo realizado para la práctica 1 realizando modificaciones específicas para esta práctica. Adicionalmente en esta fase se configuro el total de palabras, frases y trabalenguas de esta práctica, lo cual incluye audios, animaciones y textos.



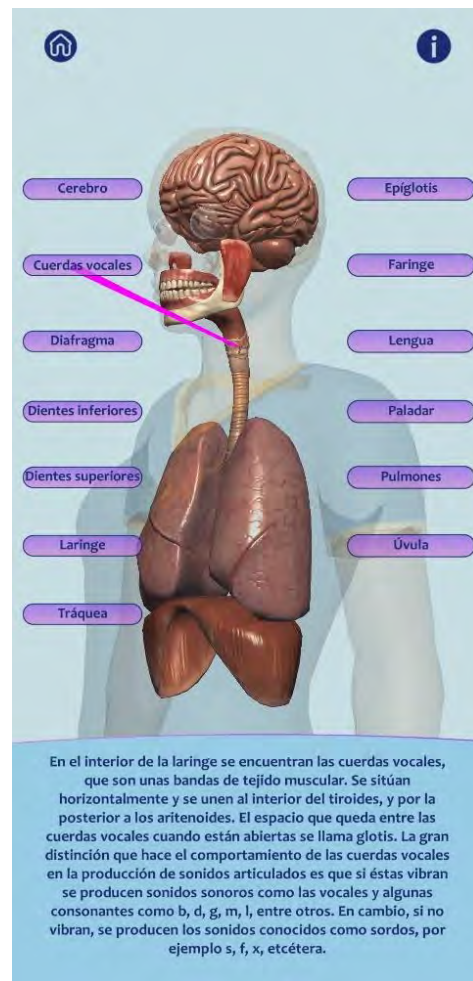
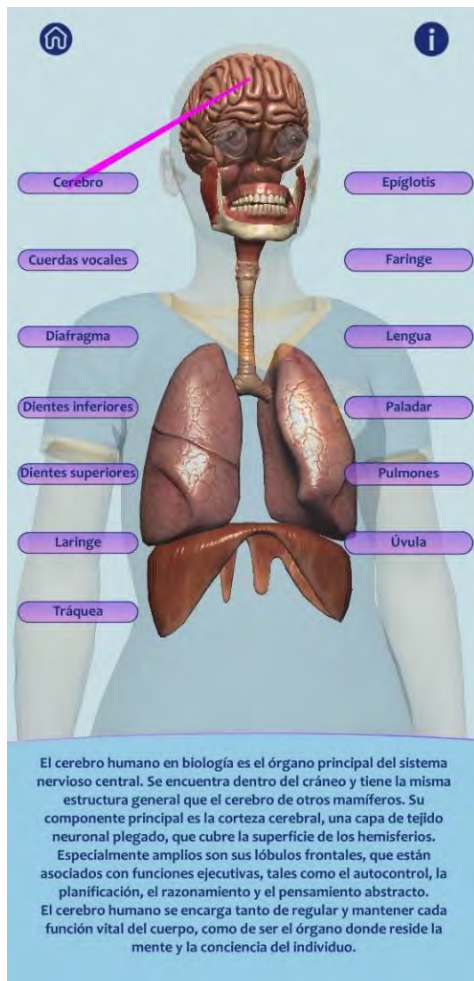
Programación de la práctica 4

De igual forma a la fase anterior, se tomó como base para la programación de la práctica 4 lo realizado para la práctica 1 realizando modificaciones específicas para esta práctica. Adicionalmente en esta fase se configuro el total de palabras, frases y trabalenguas de esta práctica, lo cual incluye audios, animaciones y textos.



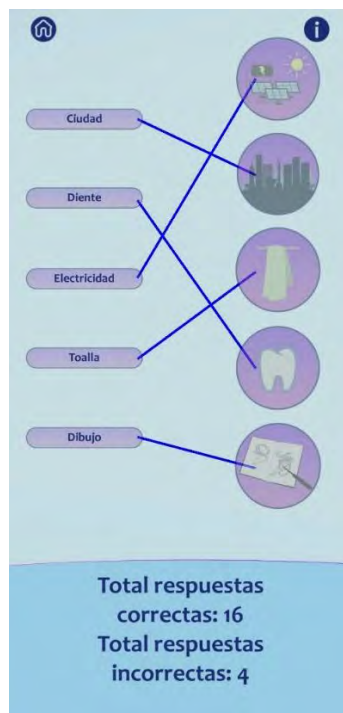
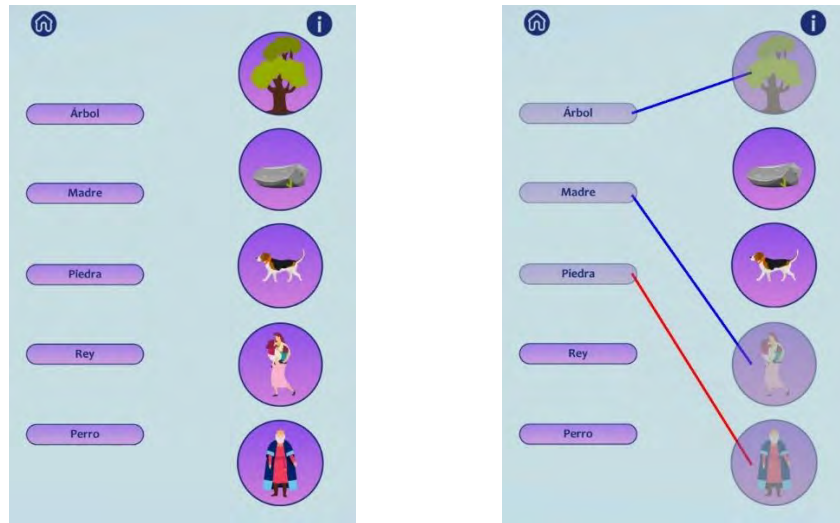
Programación de la sección de órganos del AFH

Se creo la escena para mostrar los órganos del aparato fonador humano, se integro los textos proporcionados por las responsables del proyecto, se programó la interacción de traslación y rotación del modelo y se programo la interacción para mostrar los textos y la línea de indicación de ubicación de cada órgano.



### Programación del juego 1

Se programo la dinámica del primer juego, consiste en relación de dos columnas compuestas por palabras del lado izquierdo e imágenes del lado derecho, el número de elementos en ambos lados es el mismo por lo que al realizar la relación no sobran elementos. El audio correspondiente a la palabra se reproduce al oprimir los elementos del lado izquierdo. Al finalizar el juego la aplicación indica el número de respuestas correctas e incorrectas.



### Programación del juego 2

Se programo la dinámica del segundo juego, consiste en relación de dos columnas compuestas por frases incompletas del lado izquierdo y palabras del lado derecho, el numero de elementos en la columna de la derecha es del doble que la columna del lado izquierdo por lo que al realizar la relación sobran elementos del lado derecho. El audio correspondiente a la palabra se reproduce al oprimir los elementos del lado izquierdo. Al finalizar el juego la aplicación indica el número de respuestas correctas e incorrectas.



Programación del examen

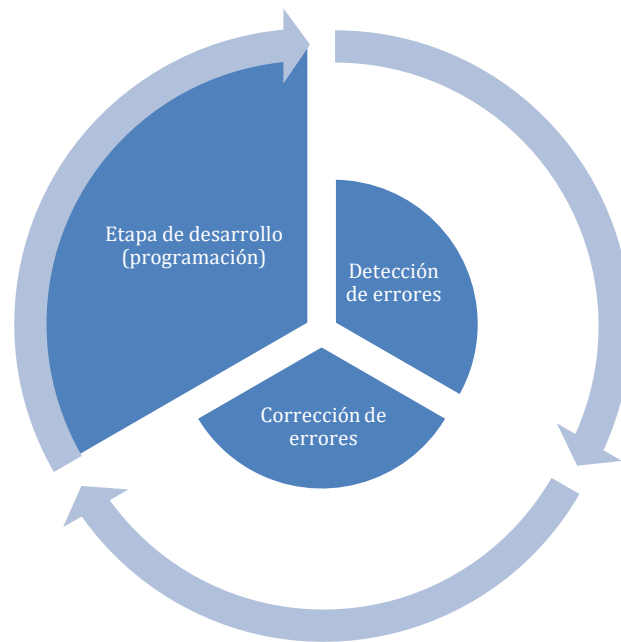
Dada su gran similitud con las prácticas, se tomó como caso para la programación el código de la práctica 1 realizando las modificaciones específicas para el examen, a diferencia de las prácticas, en el examen no es posible avanzar sin haber respondido. Se configuro el total de palabras y frases correspondientes, lo cual incluye audios, animaciones y textos.





## Corrección de errores

Al finalizar el desarrollo se llevó a cabo una serie de pruebas para detectar errores en el funcionamiento de la aplicación. En realidad, este proceso se realizó de forma cíclica al finalizar cada una de las etapas de desarrollo referentes a programación con el fin de realizar las correcciones más evidentes de forma inmediata, con la diferencia de que en la última fase del desarrollo estas pruebas se realizaron de forma más exhaustiva.





---

## PRUEBAS CERRADAS CON UN GRUPO DE USUARIOS

La aplicación tuvo un periodo de pruebas cerradas con un grupo usuarios elegidos por las responsables del proyecto con el fin de validar la facilidad de uso, detectar errores y validar los contenidos de la aplicación.

Para realizar las pruebas cerradas se aprovecharon las plataformas de publicación de Android y iOS que permiten realizar pruebas alfa y beta antes de lanzar la publicación final. Las responsables del proyecto se encargaron de seleccionar al grupo para la realización de pruebas y los correos de estos usuarios fueron ingresados en las plataformas para darles acceso a la versión anticipada de la aplicación.

Para instalar la aplicación de prueba en dispositivos iOS los pasos eran los siguientes:

- Instalar TestFlight en el dispositivo iOS que usará para las pruebas. Puede darle clic al siguiente enlace desde su dispositivo iOS o buscarla por nombre directamente en la AppStore de su dispositivo. <https://apps.apple.com/es/app/testflight/id899247664>
- Abrir el correo electrónico de invitación o abrir el siguiente enlace en el dispositivo iOS. <https://testflight.apple.com/join/JchyekOY>
- Tocar o dar clic en "Ver en TestFlight" o "Empezar las pruebas"; o tocar o dar clic en Instalar o Actualizar para la aplicación Pronuncia español con Sarahi.

Para instalar la aplicación de prueba en dispositivos Android los pasos eran los siguientes:

- Abrir el correo electrónico de invitación o abrir el siguiente enlace en el dispositivo Android. <https://play.google.com/apps/internaltest/4698604742904004571>
- Tocar o dar clic en "Unirse al programa" y posteriormente en "descárgala desde Google Play".
- Una vez abierta la tienda Google Play, tocar o dar clic en "Instalar".

Al finalizar las pruebas con el grupo de usuarios se llevó a cabo una evaluación por parte de las responsables del proyecto para recabar las opiniones y observaciones de los usuarios de prueba, se realizó la corrección de los errores detectados y se realizaron pequeñas mejoras y cambios menores tomando en cuenta las opiniones.

## LANZAMIENTO DE LA APLICACIÓN EN LAS TIENDAS

El enlace de la publicación en las tiendas de aplicaciones es:

- <https://play.google.com/store/apps/details?id=mx.unam.dgtic.dvr.v.espanolsarahi> (Android)
- <https://apps.apple.com/us/app/pronuncia-espa%C3%B1ol-con-sarahi/id1478639289> (iOS)

Requisitos para la instalación de la aplicación

- Celular o tableta con Android 4.1 o superior (Android)
- Celular o tableta con iOS 9.0 o superior (iOS)