

Galileo
UNIVERSIDAD
La Revolución en la Educación

Ingeniería en Sistemas, Informática y
Ciencias de la Computación

Seminario Profesional II

MAYALENG



Autores:
Douglas Figueroa
Alexander Baquiaux

Supervisado por:
Ing. Jack Trachtenberg
Ing. Axel Benavides

Índice

1. Introducción	4
2. Objetivos	5
2.1. Objetivo General	5
2.2. Objetivos Específicos	5
3. El Problema	6
3.1. ¿Cuál es el problema?	6
4. Estudio de Factibilidad	7
4.1. Factibilidad Funcional	7
4.1.1. Resultados de Encuesta	7
4.1.2. Tablas Comparativas	10
4.1.3. Conclusión	10
4.2. Factibilidad Técnica	10
4.2.1. Conclusión	11
4.3. Factibilidad Económica	11
4.3.1. Conclusión	12
5. Marco Teórico	13
5.1. Lenguas Mayas	13
6. MayaLeng	17
6.1. Diagrama Entidad-Relación	18
6.2. Diagrama de Arquitectura	19
6.3. Diagrama de Arquitectura con Tecnologías	20
6.4. Diagrama de Secuencia - Palabra	20
6.5. Diagrama de Secuencia - Oraciones	21
6.6. Diagrama de Casos de Uso	21
6.7. Algoritmos implementados	22
6.7.1. Traducción de español a kaqchikel	22
7. Conclusiones	23
8. Recomendaciones	24
9. Referencias	25
10. Anexo	26
10.1. Calendarización	26
10.2. Encuesta	26
11. Glosario	27

Índice de figuras

1.	Aprender una lengua maya	7
2.	Comunicación en lengua maya	7
3.	Cultura importante	8
4.	Una aplicación móvil	8
5.	Pago por la aplicación	9
6.	Inversión de tiempo	9
7.	Mapa de lenguas mayas	13
8.	Tabla de lenguas mayas	14
9.	Porcentaje de hablantes de lenguas mayas	15
10.	Oración español-kaqchikel	18
11.	Diagrama Entidad Relación	18
12.	Diagrama de Arquitectura	19
13.	Diagrama de Arquitectura Logos	20
14.	Diagrama de Secuencia	20
15.	Diagrama de Secuencia de Oraciones	21
16.	Diagrama de Caso de Uso	21
17.	Diagrama de Gantt	26

Índice de tablas

1.	Comparativa de Funcionalidades	10
2.	Comparativa de Características	10
3.	Factibilidad Económica	12
4.	Tecnologías Involucradas	15

1. Introducción

Mayaleng más que una aplicación es un algoritmo con la capacidad de recibir un input, en nuestro caso, una palabra, una frase, un párrafo, una cantidad de texto muy grande, en español y retornar un output, (el texto de entrada), en una lengua maya.

Una herramienta de traducción y para conocimiento de la cultura de Guatemala.

La cultura en Guatemala es muy grande, uno de los propósitos de este proyecto es poder aportar a expandir el conocimiento en las lenguas mayas, una rama que forma gran parte de la cultura, por medio de un algoritmo de traducción, utilizando diferentes tecnologías, hacemos dicho aporte, creando un traductor de español e inglés a una lengua maya, el Kaqchikel.

Una aplicación móvil y web nos ayudará a expandir el conocimiento de esta lengua maya, la tecnología cada vez es más accesible, los teléfonos inteligentes y los computadores cada vez son más comunes en los hogares, lo que nos ayuda a deducir que el alcance puede ser mayor. Pero no solamente estamos enfocando esta herramienta al uso particular de personas, también se enfoca a que sea una herramienta de trabajo para quienes enseñan a otros el español, en este caso es un profesor quien utiliza la herramienta y es cierto que esta persona ya conoce la lengua maya pero con la herramienta puede generar su propia documentación para los alumnos que este tenga.

Utilizamos muchas herramientas para la construcción de este proyecto, MySQL, LiguaKit, PHP, Symfony, Ionic, API's; herramientas de las cuales iremos hablando en el desarrollo del proyecto.

2. Objetivos

2.1. Objetivo General

Crear una herramienta de traducción de la lengua maya Kaqchikel.

2.2. Objetivos Específicos

- Crear una aplicación móvil (Android y iOS) para traducción del Kaqchikel.
- Dar a conocer la cultura de nuestro país Guatemala.
- Brindar la capacidad a profesores y personas en general de traducir documentos para facilitar su trabajo en los lugares dónde imparten clases dentro de nuestro país y los cuales no dominan el español.

3. El Problema

3.1. ¿Cuál es el problema?

El problema analizado para querer desarrollar esta herramienta fue desde un punto de vista cultural, ya que en Guatemala existen muchas lenguas mayas, de las cuales pocas personas en el país dominan como mínimo una. Estas forman parte de nuestro patrimonio cultural. Alrededor del mundo Guatemala es conocida por los puntos turísticos con los que contamos, regularmente estos se encuentran en los departamentos donde al menos existe una lengua Maya. Como guatemaltecos deberíamos de poder comunicarnos con la gente de nuestro país, ya que en algunos departamentos no se habla mucho en español sino que hablan en una lengua maya, sin embargo ese no es el caso. No hemos encontrado tantas herramientas claras que nos faciliten dicha tarea, queremos ser de los primeros en brindar dicha herramienta.

En este momento sugerimos Kaqchikel, la forma en que solucionamos el problema de la comunicación es teniendo una aplicación móvil la cuál hará una traducción en nuestro idioma natal a Kaqchikel, sencillamente quien utilice la aplicación deberá ingresar el texto y en una casilla aparte aparecerá la traducción del texto, algo similar a google translate.

Le ahorraremos al usuario tener que escribir algunas de las frases o palabras más utilizadas, ya que tendremos una sección donde diremos como se dicen esas frases, cosas básicas como saludar, despedirse, hasta algo más formal, un ejemplo sería preguntar dónde se encuentra el bano, preguntar el nombre de una persona y frases similares.

4. Estudio de Factibilidad

4.1. Factibilidad Funcional

4.1.1. Resultados de Encuesta

1. ¿Le gustaría aprender una lengua Maya? (79 respuestas)

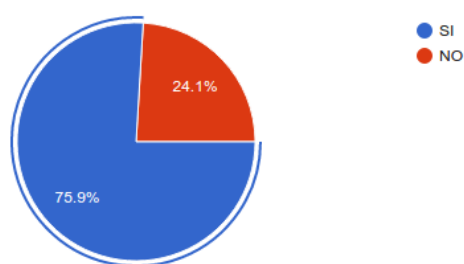


Figura 1: Aprender una lengua maya

Del total de personas que contestaron la encuesta podemos observar que poco más del 75 % de las personas está interesada en aprender una lengua maya.

2. ¿Al visitar un departamento en Guatemala le gustaría poder comunicarse con personas que solamente conocen una lengua Maya?

(79 respuestas)

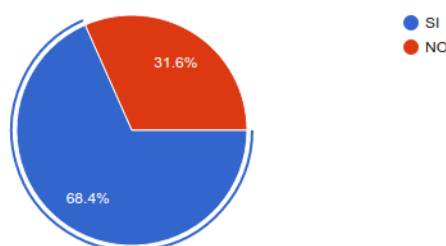


Figura 2: Comunicación en lengua maya

En la gráfica podemos observar que casi la tercera parte de las personas les gustaría entablar una conversación en lengua maya, estas personas pueden estar enlazadas a la pregunta 1, que contestaron con un sí.

3. ¿Cree que aprender sobre la cultura de nuestro país es importante?

(79 respuestas)

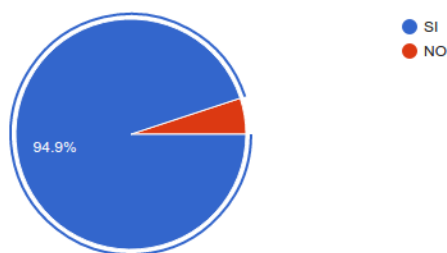


Figura 3: Cultura importante

En esta pregunta vemos que casi todos contestaron que sí, hay un pequeño porcentaje que piensa que la cultura de nuestro país no es algo importante.

4. ¿Usaría una aplicación móvil para aprender cosas básicas sobre lenguas mayas?

(79 respuestas)

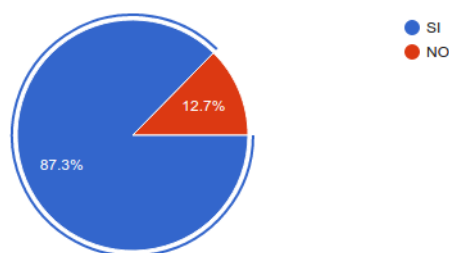


Figura 4: Una aplicación móvil

Todas las personas que entrevistamos cuentan con un smartphone, de todas esas personas la mayoría está dispuesta a instalar y utilizar nuestra aplicación en su teléfono.

5. ¿Pagaría por la aplicación? (79 respuestas)

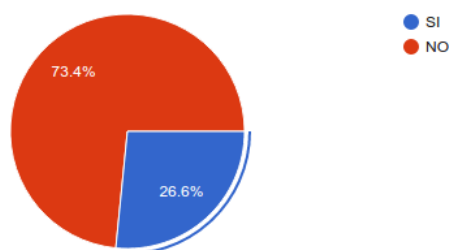


Figura 5: Pago por la aplicación

En esta pregunta creímos que el 100 % de las personas contestaría con un no, sin embargo un pequeño porcentaje de personas estaría dispuesta a pagar por la aplicación.

6. ¿Estaría dispuesto a invertir 10 minutos al día para conocer una lengua Maya?

(79 respuestas)

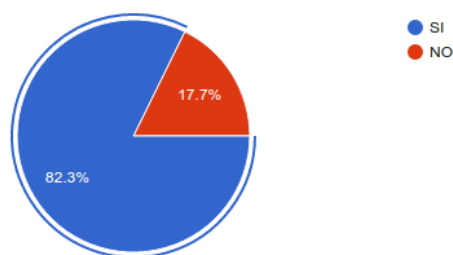


Figura 6: Inversión de tiempo

Diez minutos es un pequeño tiempo para revisar una aplicación móvil, tiempo que la mayoría de los entrevistados estarían dispuestos a utilizar para revisar el traductor maya.

4.1.2. Tablas Comparativas

Funcionalidad	Google Translate	Glosbe.com	Profesor	MayaLeng
Traducción de palabras	✓	✓	✓	✓
Traducción bidireccional	✓	x	✓	x
Traducción de frases	✓	x	✓	✓
Traducción precisa	✓	x	✓	✓
Herramienta de traducción de documentos	✓	x	✓	✓

Tabla 1: Comparativa de Funcionalidades

A nivel de funcionalidad nuestro competidor más grande es un profesor, quién posee las habilidades necesarias para desempeñar el trabajo de traductor, sin embargo hablamos de una persona contra una herramienta de tecnología que es más factible.

Característica	Google Translate	Glosbe.com	Profesor	MayaLeng
Descripción de palabras	x	✓	✓	x
Intuitivo	✓	x	✓	✓
Es gratis	✓	✓	x	✓
Siempre disponible	✓	✓	x	✓
Fácil de obtener/contactar	✓	✓	x	✓

Tabla 2: Comparativa de Características

Cuando evaluamos características la tecnología toma el control en aportar más cosas factibles que un profesor (nuestra mayor competencia), y podemos observar que MayaLeng cumple con casi todas las características de la evaluación realizada.

4.1.3. Conclusión

Nuestra herramienta cumple con ser una herramienta factible, en base a las encuestas, conocemos la opinión de las personas y vemos que comportamiento tendrían si tuviesen la herramienta en sus manos, con las tablas comparativas pudimos ver a que nivel nos encontramos comparándonos con otras soluciones similares.

4.2. Factibilidad Técnica

Pretendemos aprovechar el apogeo de los móviles y tratar de crear una buena oportunidad para introducir nuestro proyecto.

Actualmente existen dos grandes sistemas operativos que domina la industria de los móviles: iOS y Android. Estamos conscientes de que desarrollar de forma nativa para ambas plataformas nos representaría un poco más de tiempo, que implicaría restarle tiempo a la parte que en verdad es importante. Por lo cual usaremos IONIC, framework que nos permite desarrollar

de manera sencilla aplicaciones móviles usando las tendencias de Responsive Web Design.

Lo interesante de este proyecto no radica en la aplicación móvil, de hecho la aplicación sólo será una forma de consumir nuestro verdadero sistema.

La idea de este proyecto radica en hacer un compilador que pueda usar una gramática y una fuente de palabras de esa gramática, y traducir de ellas las palabras al español. En pocas palabras nuestro proyecto radica en hacer un compilador de idiomas mayas.

Durante el transcurso de nuestra carrera ya hicimos un compilador, con todas las fases básicas que uno de ellos debe tener. Ahora nos pusimos el reto de hacer un compilador genérico. Para esta primera versión usaremos dos idiomas Mayas.

Creemos tener los conocimientos necesarios para la construcción de esto. Uno de los inconvenientes más grandes era que ninguno de los dos sabemos hablar un idioma Maya, pero nos apoyamos en la documentación que distintas personas a lo largo de la historia han construido.

La idea es construir un API al que se le pueda dar como input un texto, párrafo, documento y que la salida sea otro documento pero con texto traducido. Ahora tenemos una base de datos con 15 mil palabras aproximadamente del Kaqchikel. La DB esta montada en un DBMS MySQL. El API fue construido en Symfony2.

El core de traducción será hecho en Java, usando herramientas de parseo y análisis sintáctico como Flex.

4.2.1. Conclusión

La tecnología utilizada cubre con las necesidades para el proyecto, nos ayuda a que las tareas sean optimizadas, utilizando menos recursos y optimización de tiempo, lo cuál es un punto muy importante a considerar.

4.3. Factibilidad Económica

Basándonos en nuestra calendarización de tareas para el desarrollo del proyecto, hemos estimado los siguientes costos para el desarrollo de la aplicación, los gastos son a nivel de costo de servidor, por compra del dominio, licencias para subir aplicación y gastos en papel, en impresión del manual de usuario.

Tarea	Días Trabajados	Costo (Q.)
Configurar servicios iniciales. Montar la base de datos del proyecto anterior.	7	70
Generar la documentación adecuada del algoritmo de traducción. Extraer las reglas de la gramática de libros y textos.	7	130
Lanzar primera versión del API. Traducción de oraciones sencillas. App consumiendo API.	7	192
Documentación final y manual de usuario. Lanzamiento final.	7	50
Total		442

Tabla 3: Factibilidad Económica

4.3.1. Conclusión

El costo monetario del desarrollo de nuestra herramienta no es tan alto debido a que contamos con las herramientas para prueba (teléfono con Android y iPad) y pocas cosas son en las que hemos invertido.

5. Marco Teórico

5.1. Lenguas Mayas

En Guatemala contamos actualmente con más de 20 lenguas mayas, los Acuerdos de Paz firmados en diciembre de 1996 hacen un compromiso de estado, el reconocimiento de los diferentes idiomas del país, lo cual hace que el país sea reconocido como uno multilingüe, y se hace constar en la Constitución que los idiomas mayas deberán respetarse y difundirse. Se han hecho esfuerzos, sin embargo los pocos habitantes que quedan hacen difícil la tarea. Muchos jóvenes de las nuevas generaciones no llegan a aprender el idioma indígena de sus padres. Actualmente los idiomas de mayor habla son el Kekchí, el Quiché, el Kaqchikel, el Mam y el Tzutujil, los cuales tienen algunos vocablos y reglas gramaticales en común.

Las lenguas mayenses derivan del protomaya, una protolengua (reconstrucción probable de la lengua origen de un grupo de lenguas), que pudo haberse hablado hace unos 5000 años a juzgar por el grado de diversificación interna en una región cercana a donde actualmente se hablan lenguas mayenses.

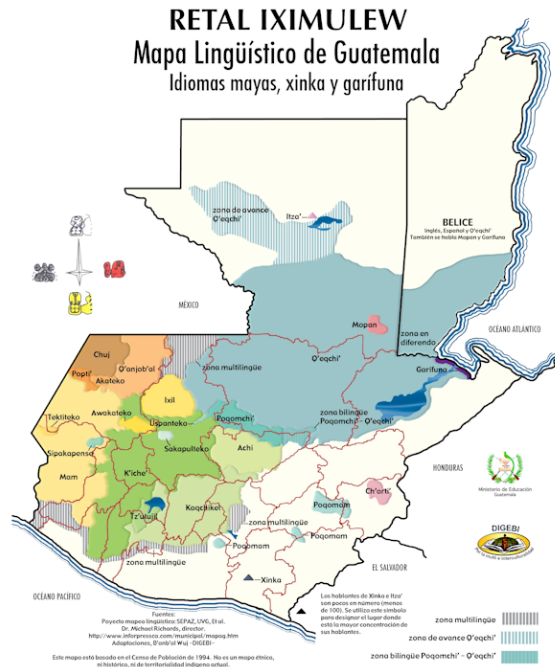


Figura 7: Mapa de lenguas mayas
 Las lenguas mayas están distribuidas por todo el país.

Idioma ▲	Familia ◆	Rama ◆	Hablantes maternos ◆
Acateco	Maya	Q'anjob'al	48.500
Achí	Maya	Kiche'	85.552
Aguateco	Maya	Mam	18.000
Chortí	Maya	Chol	30.000
Chuj	Maya	Q'anjob'al	40.000
Español	Indoeuropea	Latina	9.481.907
Garífuna	Arahuaca	Caribefña	5.860
Idioma xinca	Aislada	Lenguas xinca	16
Itzá	Maya	Yucateca	12
Ixil	Maya	Mam	70.000
Jakalteco	Maya	Q'anjob'al	40.000
Kaqchikel	Maya	Kiche'	500.000
Mam	Maya	Mam	480.000
Mopan	Maya	Yucateca	2.000
Pocomam	Maya	Kiche'	30.000
Poqomchi	Maya	Kiche'	92.000
Quekchí	Maya	Kiche'	555.461
Quiché'	Maya	Kiche'	1.000.000
Q'anjob'al	Maya	Q'anjob'al	77.700
Sacapulteco	Maya	Kiche'	9.763
Sipacapense	Maya	Kiche'	8.000
Tectiteco	Maya	Mam	2.265
Uspanteco	Maya	Kiche'	3.000
Zutuhil	Maya	Kiche'	88.300

Figura 8: Tabla de lenguas mayas
Cantidad de hablantes de cada lengua.

Lengua	% Hablantes
Acateco	0.3828
Achí	0.6753
Aguateco	0.1421
Chortí	0.2368
Chuj	0.3157
Español	74.8473
Garífuna	0.0463
Idioma Xinka	0.0001
Itzá	0.0001
Ixil	0.5526
Jakalteco	0.3157
Kaqchikel	3.9468
Mam	3.7890
Mopan	0.0158
Pocoman	0.2368
Poqomchi	0.7262
Quekchí	4.3846
Quiché	7.8937
Q'anjob'al	0.6133
Sacapulteco	0.0771
Sipacapense	0.0631
Tectiteco	0.0179
Uspanteco	0.0237
Zutuhil	0.6970

Figura 9: Porcentaje de hablantes de lenguas mayas

Como parte de las herramientas involucradas para el desarrollo de este proyecto tenemos:

1.	MySQL
2.	PHP
3.	RAE
4.	RAE
5.	LinguaKit
6.	Symfony
7.	Ionic
8.	Formularios de Google
9.	LaTeX
10.	Wordpress

Tabla 4: Tecnologías Involucradas

Habíamos mencionado en la sección 3, ***El Problema***, que nuestra problemática era desde un punto cultural, un problema de comunicación interna por falta de conocimiento, porque no conocemos las lenguas mayas, no las podemos hablar, y tampoco encontramos información fácil sobre ellas, no es cómo hacer una traducción en Google translate y saber como se escribe alguna palabra que necesitamos usar, eso nos lleva al siguiente análisis.

6. MayaLeng

¿En qué consiste?

La solución planteada a esta problemática es la construcción de un algoritmo que nos ayude a establecer esta comunicación que no se posee, trabajamos en un traductor con el cual podamos transformar frases en español a frases en una lengua maya, como lengua inicial hemos escogido el Kaqchikel. Una vez construido el algoritmo de traducción necesitamos un dispositivo por el cual podamos darle uso a dicha herramienta, el dispositivo que elegimos para nuestra implementación es un teléfono inteligente, lo aplicamos para teléfonos que tengan instalado los sistemas operativos **Android** y **iOS**; Android es un sistema operativo móvil creado por Google Inc, y iOS es un sistema operativo móvil creado por Apple Inc.

El mercado de teléfonos inteligentes cada vez crece más, con la llegada del 4G el mercado móvil dió un nuevo paso y más personas tienen más fácil acceso al internet por medio de un dispositivo móvil, por lo que tener aplicaciones que requieran internet cada vez es menos costoso para el usuario, en el 2014 se estimó que había una cantidad de 21.7 millones de líneas telefónicas móvil en el país.

¿Qué hicimos?

Creamos una aplicación móvil con el nombre de **MayaLeng** utilizando **ionic**, este es un framework que nos permite hacer un solo desarrollo y compilar nuestro programa una sola vez y generamos una aplicación para iOS y Android; MayaLeng consume el algoritmo de traducción.

¿Cómo funciona?

Lo que se requiere para su funcionamiento es escribir una palabra, una frase, una oración o un párrafo en una caja de texto, una vez MayaLeng haya iniciado, este texto llega al algoritmo, contamos con un diccionario de español-kaqchikel, el diccionario lo llenamos con su respectivo tipo de palabra, el cual obtuvimos utilizando la **RAE**, una institución cultural dedicada a la regularización lingüística.

Utilizamos su API con un script desarrollado en **PHP** que es un lenguaje de programación que se ejecuta del lado del servidor, es considerado uno de los lenguajes más flexibles, potentes y de alto rendimiento, con este script obtenemos el tipo de palabra (sustantivo, pronombre, verbo, etc) de cada una de las contenidas en el diccionario.

El diccionario está almacenado en una base de datos **MySQL** que es sistema gestor de base de datos (DBMS) de licencia gratuita; mandamos nuestra oración a otra herramienta llamada **LinguaKit** que nos devuelve cada palabra de nuestra oración con su respectivo tipo de palabra. Las palabras

podrían ser artículo, pronombre, verbos; una vez tenemos esos resultados los enlazamos con el tipo de palabra con el que ya contamos en la base de datos y ya teniendo la información, con **Symfony**, un framework diseñado para optimizar el desarrollo de las aplicaciones web basado en el patrón MVC (Modelo Vista Controlador), construimos la oración.

¿Cómo construimos la oración?

Para explicar la construcción mostramos la siguiente imagen:

Xusík' ri ab'äj ri achi. El señor recogió la piedra.
P S

Figura 10: Oración español-kaqchikel

La estructura de una oración simple en español es:

Sujeto - Predicado

La estructura de una oración simple en Kacchikel es:

Predicado - Sujetio

Basándonos en la respuesta de *Linguakit*, que nos devuelve que palabras forman parte del sujeto, y cuales del predicado, podemos colocar cada palabra en su orden correspondiente en base a la estructura de Kaqchikel que se obtuvo, este procedimiento lo realizamos para oraciones simples.

6.1. Diagrama Entidad-Relación

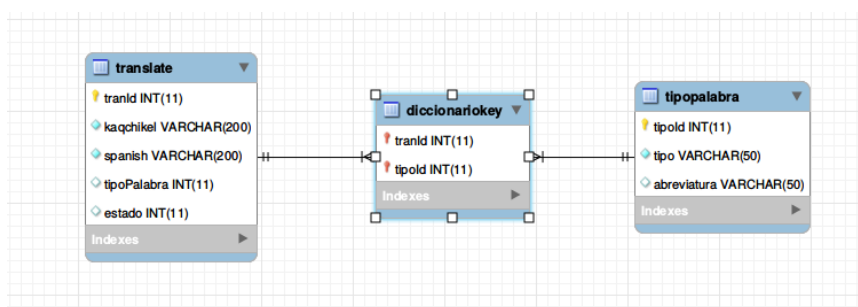


Figura 11: Diagrama Entidad Relación

La base de datos está compuesta por una tabla principal que contiene el vocabulario, con la palabra en español y su traducción a Kaqchikel, una tabla con la estructura gramatical (sujeto, verbo, adjetivo, etc), y una tercer tabla que nos servirá para enlazar nuestro vocabulario con el tipo de palabra

a la que aplique (por ejemplo: él=pronombre).

En nuestro diagrama está contemplado el caso que una palabra pudiese cumplir con más de una regla gramatical, y nos es posible almacenar esa información con nuestra tercer tabla donde tenemos las llaves primarias de cada tabla por lo cual no nos topáramos con ningún registro duplicado.

6.2. Diagrama de Arquitectura

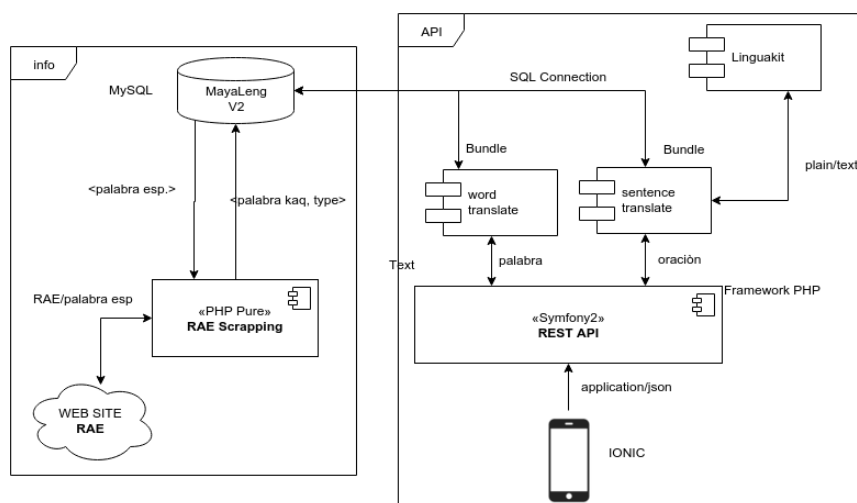


Figura 12: Diagrama de Arquitectura

Gráfico de arquitectura de MayaLeng donde explicamos cómo está construido nuestro proyecto, dividido en dos partes principales que constan de la información, que cuenta con nuestro vocabulario, la base de datos, un servicio de migración utilizando la RAE, y el API, el cual divide las oraciones y construye las nuevas en el nuevo lenguaje.

6.3. Diagrama de Arquitectura con Tecnologías

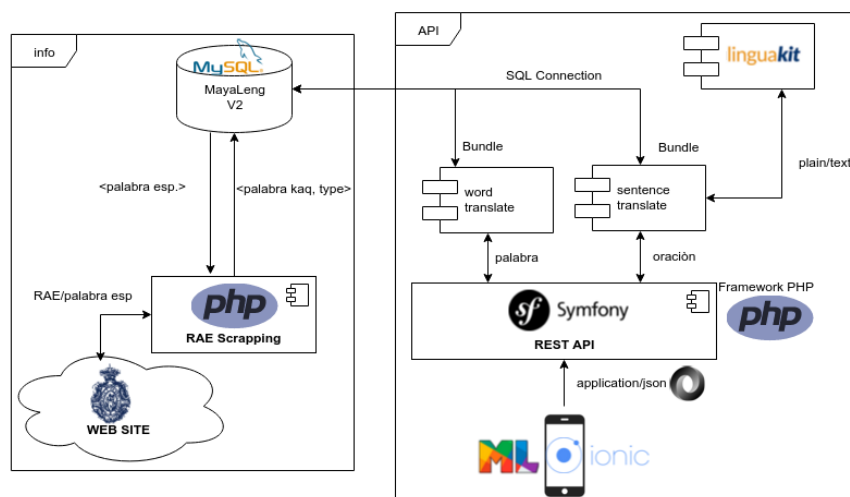


Figura 13: Diagrama de Arquitectura Logos
Tecnologías utilizadas

6.4. Diagrama de Secuencia - Palabra

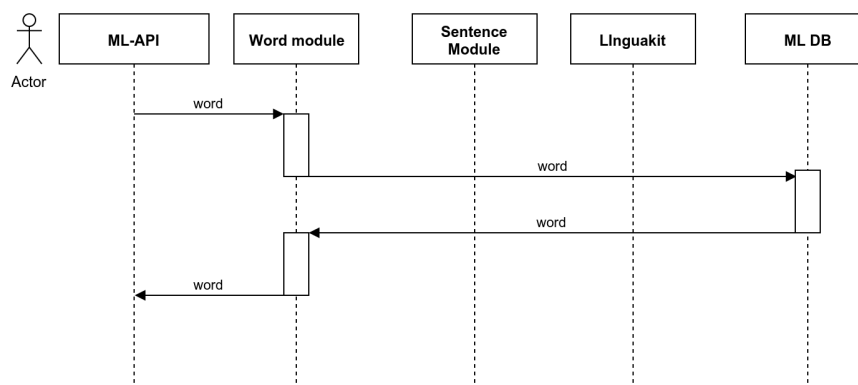


Figura 14: Diagrama de Secuencia

Aquí se presenta el flujo que tiene una palabra al ser ingresada por el usuario hasta regresar de nuevo pero ya traducida.

6.5. Diagrama de Secuencia - Oraciones

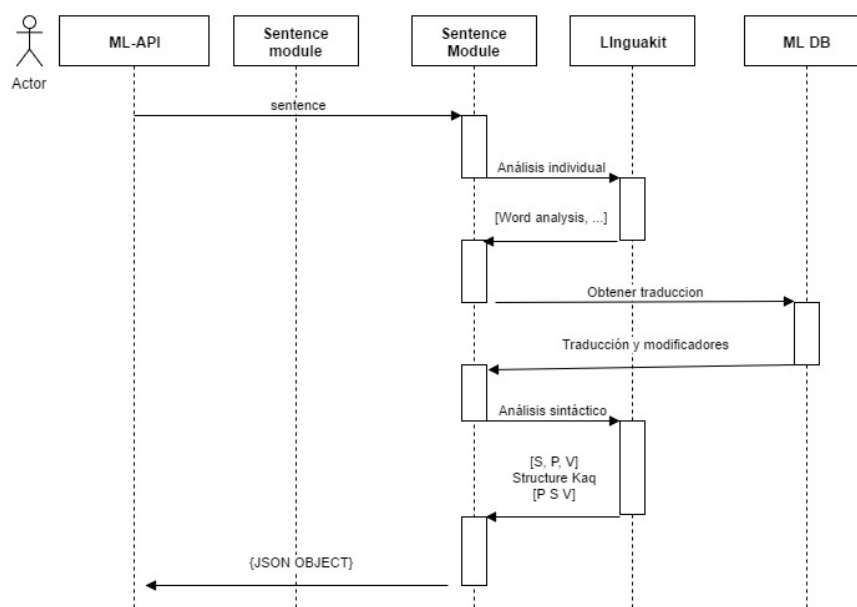


Figura 15: Diagrama de Secuencia de Oraciones

6.6. Diagrama de Casos de Uso

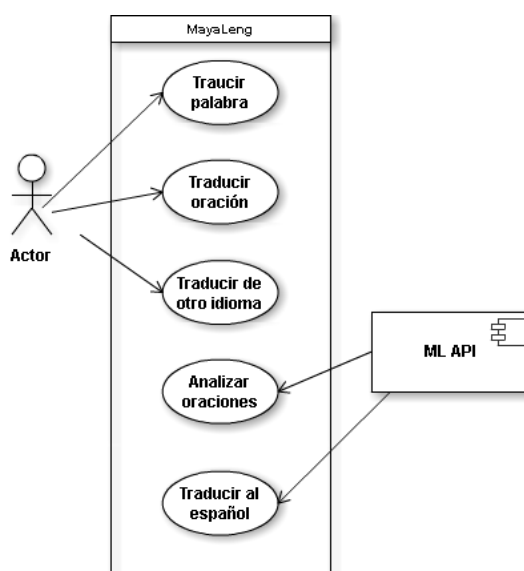


Figura 16: Diagrama de Caso de Uso

6.7. Algoritmos implementados

6.7.1. Traducción de español a kaqchikel

Lo que realiza este algoritmo, el cual es una implementación propia, se describe a continuación:

Recibe texto

Si TextoMuyLargo Entonces

 Descomponer en oraciones mas cortas

Obtener palabras de texto

Mientras Palabras > 0

 Obtener traduccion de cada palabra en kaqchikel

Analizar tipo de palabra

Analizar estructura de oracion en kaqchikel

Armar nuevo texto

Devolver nuevo texto

7. Conclusiones

- Aprendimos un poco más de la cultura de Guatemala, sobre las lenguas mayas.
- Pusimos en práctica varios temas de los aprendidos en la Universidad.
- Aprendimos nuevas tecnologías.

8. Recomendaciones

- Aprender más sobre la cultura de Guatemala.

9. Referencias

- Filiberto Patal Majzul. (2007). Ruzoltzij Ri Kaqchikel. Guatemala—: Cholsamaj.
- (2006). Lenguas mayenses. 2006, de Wikipedia Sitio web: <http://bit.ly/2dTGUVy>
- (2002). Lenguas de Guatemala. 2002, de Wikipedia Sitio web: <http://bit.ly/2fuwPpd>

10. Anexo

10.1. Calendarización

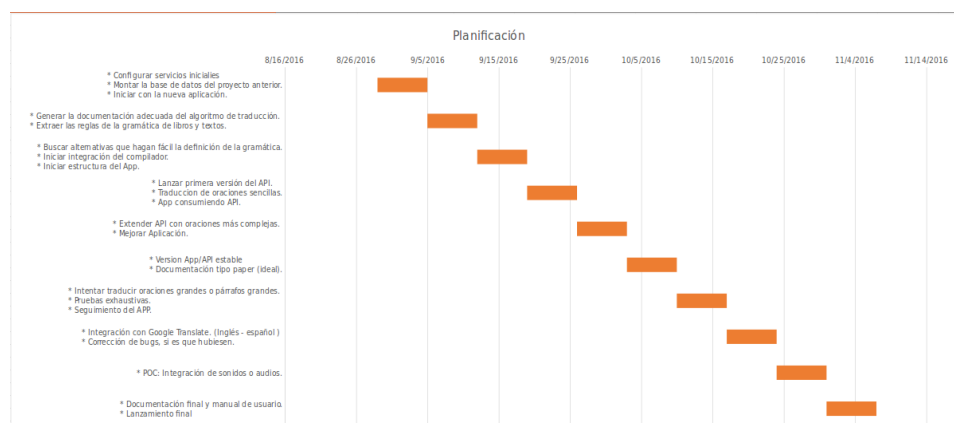


Figura 17: Diagrama de Gantt

10.2. Encuesta

- ¿Le gustaría aprender una lengua maya?
SI _____ NO _____
- ¿Al visitar un departamento en Guatemala le gustaría poder comunicarse con personas que solamente conocen una lengua maya?
SI _____ NO _____
- ¿Cree que aprender sobre la cultura de nuestro país es importante?
SI _____ NO _____
- ¿Usaría una aplicación móvil para aprender cosas básicas sobre lenguas mayas?
SI _____ NO _____
- ¿Pagaría por la aplicación?
SI _____ NO _____
- ¿Estaría dispuesto a invertir 10 minutos al día para conocer una lengua maya?
SI _____ NO _____

11. Glosario

Kaqchikel: lengua maya, parte de la familia lingüística mayense.

iOS: sistema operativo móvil de Apple

Android: sistema operativo móvil de Google

Framework: conjunto estandarizado de conceptos, prácticas y criterios para enfrentar y resolver nuevos problemas de índole similar.

Responsive Web Design: adaptar la apariencia de las páginas web al dispositivo que se esté utilizando para visualizarlas.

API: utilizada por otro software como una capa de abstracción

DB: base de datos

DBMS: programa que permite almacenar y posteriormente acceder a los datos de forma rápida y estructurada.

MySQL: sistema de gestión de bases de datos relacional

Symfony2: framework diseñado para optimizar el desarrollo de aplicaciones web