

## INDICE

	PAG.
1. INTRODUCCIÓN	03
2. PLANTEAMIENTO DEL PROBLEMA	04
3. JUSTIFICACIÓN	05
4. OBJETIVOS	05
4.1 GENERAL	05
4.2 ESPECÍFICOS	05
5. ALCANCES Y LIMITACIONES	06
6. ESTADO DEL ARTE	07
6.1. DICCIONARIO NÁHUATL-ESPAÑOL EN LÍNEA	07
6.2. PORTAL HUACHARAN	07
6.3. FREETASNLATION	08
6.4. WORDLINGO	09
6.5. TRADUCTOR ON-LINE	09
6.6. TABLA COMPARATIVA DE LOS TRADUCTORES ENCONTRADOS	10
7. MARCO TEÓRICO	11
7.1. TSOTSIL	11
7.2. MODELO DEL SICLO DE VIDA DEL SOFTWARE	13
7.3. PROGRAMACIÓN ORIENTADO A OBJETO (POO)	13
7.3.1. LOS OBJETOS	14
7.3.2. CLASES	14
7.3.3. HERENCIA	15
7.3.4. MENSAJE	15
7.3.5. ATRIBUTOS	15
7.3.6. MÉTODOS	16
7.4. TRADUCTORES	16
7.5. ENSAMBLADORES	16
7.6. PREPROCESADORES	17
7.7. LENGUAJES	17
7.8. PUNTOS PARA LA DEFINICIÓN DE UN LENGUAJE	19
7.9. PARTES DE UNA ORACIÓN	20
7.10. GRAMÁTICA	21
7.11. TIPOS DE GRAMÁTICAS JERÁRQUICAS DE CHOMSKY	21
7.12. CLASIFICACIÓN DE GRAMÁTICAS	22
7.13. GRAMÁTICA REGULAR	23
7.14. GRAMÁTICA INDEPENDIENTE DE CONTEXTO	24

<b>7.15. GRAMÁTICAS SENSIBLES AL CONTEXTO</b> .....	<b>24</b>
<b>7.16. GRAMÁTICAS NO RESTRINGIDAS</b> .....	<b>25</b>
<b>7.17. GRAMÁTICA ESPAÑOL</b> .....	<b>25</b>
<b>7.17.1. TÉRMINOS GRAMATICALES</b> .....	<b>25</b>
<b>7.17.1.1. ADJETIVO</b> .....	<b>25</b>
<b>7.17.1.2. ARTÍCULO</b> .....	<b>26</b>
<b>7.17.1.3. ADVERBIO</b> .....	<b>26</b>
<b>7.17.1.4. VERBO</b> .....	<b>28</b>
<b>8. HIPÓTESIS</b> .....	<b>29</b>
<b>9. ANÁLISIS DEL SISTEMA</b> .....	<b>30</b>
<b>9.1. GRAMÁTICA TSOTSIL</b> .....	<b>30</b>
<b>9.2. SUSTANTIVO</b> .....	<b>30</b>
<b>9.3. MACA POSESIVA</b> .....	<b>31</b>
<b>9.4. ARTÍCULO DETERMINANTE</b> .....	<b>32</b>
<b>9.5. MARCA DEL SUJETO</b> .....	<b>32</b>
<b>9.6. PREPOSICIÓN</b> .....	<b>32</b>
<b>9.7. MARCA ASPECTUAL DE TIEMPO</b> .....	<b>33</b>
<b>9.8. VERBO TRANSITIVO</b> .....	<b>34</b>
<b>9.9. VERBO INTRANSITIVO</b> .....	<b>35</b>
<b>9.10. CLÍTICO</b> .....	<b>35</b>
<b>9.11. VERBOS MULTICATEGORIA</b> .....	<b>36</b>
<b>9.12. SUFIJO DEL SUSTANTIVO</b> .....	<b>36</b>
<b>9.13. GRAMÁTICA DE ORACIONES SIMPLES EN TIEMPO PRESENTE DE</b> <b>ESPAÑOL A TSOTSIL</b> .....	<b>36</b>
<b>9.14. GRAMÁTICA TSOTSIL EN TIEMPO PRESENTE</b> .....	<b>37</b>
<b>9.15. GRAMÁTICA TSOTSIL EN TIEMPO PASADO</b> .....	<b>38</b>
<b>9.16. GRAMÁTICA ESPAÑOL ADAPTADA AL TSOTSIL</b> <b>CON ORACIONES SIMPLES EN PRESENTE Y PASADO</b> .....	<b>39</b>
<b>9.17. EJEMPLOS DE ORACIONES SIMPLES EN TSOTSIL</b> .....	<b>41</b>
<b>9.18. DIAGRAMAS DE CASOS DE USO</b> .....	<b>46</b>
<b>9.18.1 DESCRIPCIÓN</b> .....	<b>47</b>
<b>9.19. DISEÑO DEL SISTEMA</b> .....	<b>49</b>
<b>9.19.1. DIAGRAMA DE CLASES</b> .....	<b>49</b>
<b>9.19.2. ESTRUCTURA DEL LEXICÓN</b> .....	<b>51</b>
<b>10. DESARROLLO DEL SISTEMA</b> .....	<b>52</b>
<b>10.1. DIAGRAMA DE BLOQUES</b> .....	<b>54</b>
<b>10.2. DIAGRAMA DE SECUENCIA</b> .....	<b>55</b>
<b>11. CASOS DE PRUEBA</b> .....	<b>56</b>
<b>12. CONCLUSIONES Y TRABAJOS FUTUROS</b> .....	<b>59</b>
<b>13. BIBLIOGRAFÍA</b> .....	<b>60</b>

## 1. INTRODUCCIÓN

En este proyecto se tiene como objetivo desarrollar un sistema traductor del **Español a la lengua Tsotsil**, principalmente es una herramienta que traducirá oraciones simples de tiempo presente, pasado y futuro, con el fin de apoyar a los alumnos hablantes de tsotsil también será de mucho apoyo para las instituciones que se dedican al rescate, enriquecimiento y fortalecimiento de las lenguas indígenas, promoviendo la conservación de la lengua.

La Secretaria de Educación Pública (SEP) y el Instituto Nacional de Estadística y Geografía (INEGI) establecen que Chiapas es uno de los principales estados del país con alto índice en rezago educativo. Este problema en el estado tiene su mayor auge en las poblaciones indígenas.

Se da a conocer la problemática que se pretende resolver en el presente proyecto, los objetivos que se desean alcanzar, así también la justificación, se analizará la información necesaria que será útil para realizar el proyecto, la técnica que se aplicará para resolver el problema, así como parte de la gramática de la lengua tsotsil que será la base para el desarrollo del proyecto.

## **2. PLANTEAMIENTO DEL PROBLEMA**

El desarrollo de esta herramienta podría beneficiar a las personas y a la educación de las poblaciones indígenas hablantes de la lengua tsotsil, debido a que no cuentan con suficientes materiales, herramientas didácticas para poder tener una mejor educación y un conocimiento adecuado de las cosas, sucesos o avances tecnológicos.

Ya que por ese motivo la mayoría de municipios del estado de Chiapas tienen un gran índice de analfabetismo ya que no cuenta con la educación adecuada y también por que no cuentan con suficientes materiales traducidos a la lengua tsotsil para poder tener una mayor conocimiento y enriquecer su vocabulario y sus conocimientos.

### **3. JUSTIFICACIÓN**

El desarrollo del sistema traductor contribuirá a la preservación de la lengua ya que Con ello se pretende que la educación indígena sea de mejor aprovechamiento y así como también los niños como los adultos conserven sus valores y culturas, Debido a la importancia que tienen las lenguas indígenas en la actualidad.

Existen una infinidad de traductores pero ninguno estrechamente relacionado con las lenguas indígenas para el apoyo de los docentes, el tsotsil carece ampliamente de un software o una tecnología informática que permita traducir cualquier texto escrito en español a la lengua tsotsil.

Para poder conservar la lengua tsotsil es necesario tener herramientas adecuadas y una gramática definida, porque los diccionarios no son suficientes para poder interpretar las palabras por ello se desarrolla el presente proyecto.

### **4. OBJETIVOS**

#### **4.1. GENERAL**

- Traducir oraciones simples en tiempo presente y pasado escritas en español a la lengua tsotsil.

#### **4.2. ESPECÍFICOS**

1. Traducirá oraciones simples en tiempo presente del español al tsotsil.
2. Traducirá oraciones simples en tiempo pasado del español al tsotsil.
3. Traducirá oraciones simples en tiempo futuro del español al tsotsil.
4. Contará con un diccionario de datos.
5. Imprimirá el texto traducido en formato \*.pdf, \*.txt.
6. Generará un archivo de texto plano.

## 5. ALCANCES Y LIMITACIONES

El alcance que tiene el sistema traductor de español a tsotsil es que este sistema podrá realizar la traducción de las siguientes oraciones:

1. Oraciones simples en tiempo pasado
2. Oraciones simples en tiempo presente
3. Oraciones simples en tiempo futuro

Basados en la gramática del idioma español, se desarrolla una gramática formal para la lengua tsotsil donde se abordaran todos los elementos principales como son sujetos, predicados, verbos, complementos y todo esto a su vez sirve para poder hacer una traducción clara de la lengua tsotsil y así los usuarios que la usen pueden entender las cosas de mejor manera.

Las limitaciones de este sistema traductor

Una de las limitaciones que se ha encontrado en este sistema traductor es en la parte de la gramática de oraciones simples en tiempo futuro, porque aún no se tiene una formalización en la gramática.

El sistema traductor de español a tsotsil se realizo en el instituto de educación par adultos (IEA) la cual se encuentra ubicada en la 2 norte poniente #1126 Col. Centro c.p. 29000, Tuxtla Gutiérrez, Chiapas, Numero de Tel. 61 3 15 16 y 61 2 98 32.

## 6. ESTADO DEL ARTE

### 6.1 Diccionario Náhuatl-Español en línea en Aulex

Traductor de palabras en línea de la lengua náhuatl al idioma español.

Es un diccionario simplemente traduce palabras de español-náhuatl y de náhuatl-español este diccionario cuenta nada más con 60 palabras para traducir. [12]



Fig. 1.1 Diccionario Náhuatl

Traductor de palabras en línea de la lengua náhuatl al idioma español.

Es un diccionario simplemente traduce palabras de español-náhuatl y de náhuatl-español este diccionario cuenta nada más con 60 palabras para traducir y es muy obsoleto para ser un diccionario. [12]

### 6.2 Portal Huacharán

Este es un diccionario Políglota Incaico. Contiene un vocabulario políglota de 5 lenguas del país Perú. Traduce palabras en 5 lenguas del país (Cuzco, Ayacucho, Junin, Ancash, Aymara y castellano), además que se le implementa la pronunciación de cada lengua este diccionario está casi completo en el sentido de traducción de varios idiomas una desventaja es que su base de datos no cuenta con suficientes palabras para poder ser un diccionario competitivo ante los demás diccionarios la interface se muestra en la figura 1.2. [13]



Fig.1.2 Portal Huacharán

### 6.3 FREETRANSLATION

Tiene variedad en información mundial de traducciones ver fig. 1.3.

Traduce del inglés al español, inglés al italiano, inglés al portugués, inglés al ruso, inglés al noruego, inglés al chino simplificado, inglés al chino tradicional, inglés al japonés, español al inglés, francés al inglés, portugués al inglés, ruso al inglés y japonés al inglés.

Como se puede observar esta traductor está completo lo único que falta es que no tiene traducción en lenguas indígenas si no en puros idiomas como es portugués ruso por eso la importancia de hacer el sistema traductor de español-tsotsil.

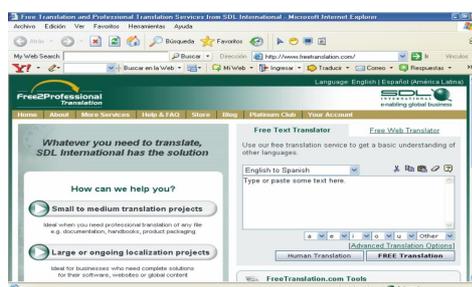


Fig.1.3freetraslation.

## 6.4 WORDLINGO

Este sistema traduce diferentes idiomas, es considera rápido en la ejecución de las tareas fig.1.4. Tiene dos opciones, una de carácter gratuito y otra pagada, que se realiza mediante la traducción profesional; el traductor gratuito tiene un límite de 150 palabras para la caja de diálogo. *WorldLingo* te permite acceder a un sitio para copiar e imprimir partes con propósitos de información y referencia. [14]



Fig.1.4 Wordlingo.

## 6.5 Traductor On-line:

Permite traducir gratis páginas Web y texto fig.1.5 Traduce una gama de lenguas de varios países del mundo. Español, inglés, alemán, francés, italiano, japonés y portugués. [15]



Fig1.5 Traductor On-line

## 6.6 Tabla de comparación entre los traductores encontrados

IDIOMA	TRADUCTORES							DICIONARIOS		
	Free Translation	Traductor Diidx	Portal Huascarán	Traductor WorldLingo	Traductor On-line	Traductor Im Translat	BableFish	FreeLan g	Diccionario SEF	Diccionario AULEX
Inglés	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Francés	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Alemán	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Chino	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Portugués	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Italiano	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Griego	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Japónés	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗
Inca	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗
Nahuatl	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓
Zapoteco	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗
Tsotsil	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗

Tabla 1 tabla comparativa.

## 7. MARCO TEÓRICO

### 7.1 Tsotsil

El Tsotsil se habla en 18 municipios del estado de Chiapas, que se encuentra al noroeste y suroeste de la ciudad de San Cristóbal de las casas Chiapas. El tsotsil se habla en las comunidades de El Bosque, Chalchihuitán, Chamula, Chenalhó, Huixtán, Mitontic, Zinacantán, Larráinzar, Pantelhó, Huitiupán, Simojovel de Allende, Totolapa, Jitotol, Pueblo Nuevo Solistahuacán, Bochil, Ixhuatán, Soyaló, Venustiano Carranza, Ixtapa, Teopisca y San Cristóbal de Las Casas. En Chiapas, el 26% de la población de 5 y más años hablan alguna lengua indígena, 73.7% se comunica en español, Chiapas tiene el segundo grupo más numeroso de indígenas. De acuerdo con los datos proporcionados por el Instituto Nacional de Estadística, Geográfica e Informática en 2000 había 291 550 fig. 1.6.

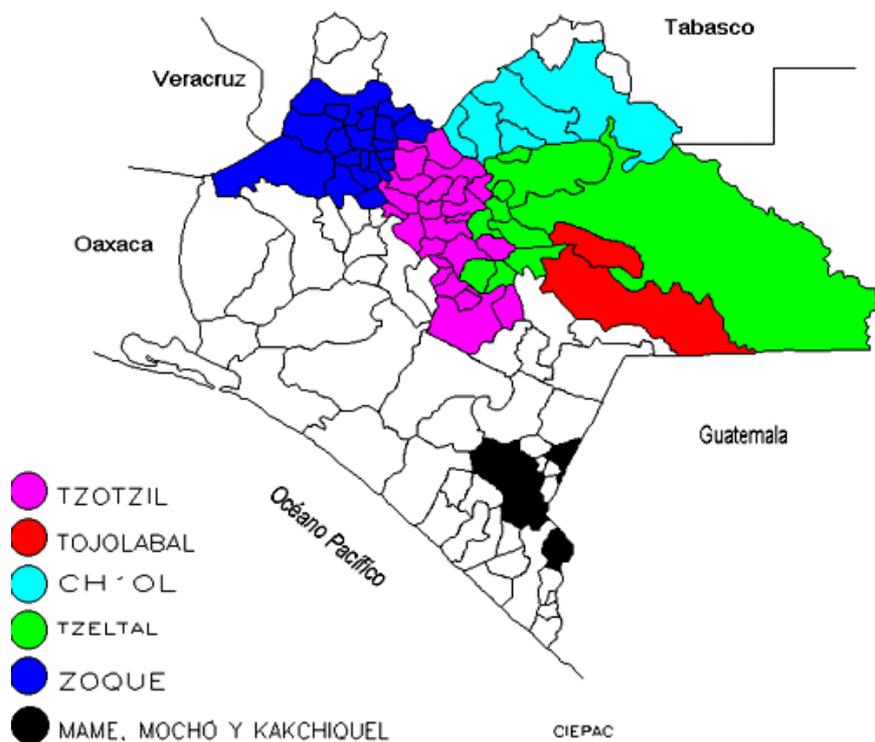


Fig. 1.6 mapa de Chiapas.

En el año 2000 la proporción de población que hablaba lengua indígena y española fue de 61.2 por ciento.

En Chiapas, en el año 2005, poco más de 957 mil personas de 5 y más años hablan alguna lengua indígena. De éstos, 73.6% se comunica también en español;

Las lenguas indígenas con mayor número de hablantes son: el Tzeltal, Tzotzil y Chol que agrupan 88.3% del total de la población hablante.

La mayoría de la población Tsotsil se localiza en la zona Altos. A pesar de la emigración y del transcurrir del tiempo, conservan la memoria cultural de sus antepasados en sus prácticas sociales y en la lengua materna, que es fundamental en todas las relaciones humanas y en la comunicación de nuestros pueblos.

Lenguas	Población indígena total	% de la pob. Indígena total
tseltal	362,658	37.9
tsotsil	320,921	33.5
chol	161,794	16.9
zoque	43,936	4.6
tojolabal	42,798	4.5
Resto de las lenguas	25,118	2.6

Tabla 2 Porcentajes de indígenas.

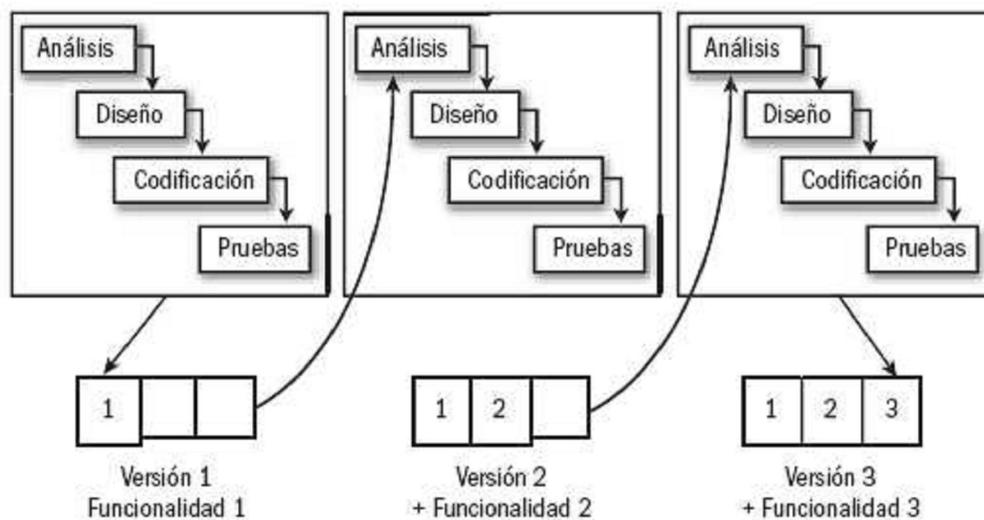
En Chiapas a igual que en todo el país el idioma oficial es el español. Sin embargo, también se hablan varias lenguas autóctonas, que son producto de la enorme herencia cultural que aun guardan nuestros pueblos. Estas lenguas son idiomas, no dialectos como comúnmente se les llama, ya que cuentan con todas las características gramaticales y fonéticas, como cualquier otro idioma, como el inglés, el francés o el alemán.

## 7.2 Modelo del ciclo de vida del software.

### METODO ITERATIVO INCREMENTAL.

Este modelo de ciclo de vida se basa en la filosofía de construir incrementando las funcionalidades del programa.

Se realiza construyendo gradualmente las capacidades del software.



**Fig.1.7 Método iterativo incremental.**

Se suele utilizar en los que los requerimientos que no están claros de parte del usuario, por lo que se hace necesario la creación de distintos prototipos para presentarlos y conseguir la conformidad del cliente.

## 7.3 Programación orientada a objetos (POO).

La programación orientada a objetos es un importante conjunto de técnicas que pueden utilizarse para hacer el desarrollo de programas más eficientes, a la par que mejora la fiabilidad de los programas de computadora. En la programación orientada a objetos, los objetos son los elementos principales de construcción.

“un método de implementación en el que los programas se organizan como colección cooperativas de objetos, cada uno de los cuales representan una instancia de alguna clase, y cuyas clases son todos todas miembros de una jerarquía de clases unidas mediante relaciones de herencia”

Los conceptos fundamentales de programación son:

- Objetos
- Clases
- Herencia
- Mensajes
- Polimorfismo

### **7.3.1 Los objetos**

La idea fundamental en los lenguajes orientados a objetos es combinar en una sola unidad datos y funciones que operan sobre esos datos. Tal unidad se denomina objeto. Por consiguiente, dentro de los objetos residen los datos de los lenguajes de programación tradicional, tales como números, arrays, cadenas y registros, así como funciones o subrutinas que operan sobre ellos.

Las funciones dentro del objeto son el único medio para acceder a los datos privados de un objeto. Si se desea leer un elemento datos de un objeto se llama a la función miembro del objeto. Se lee el elemento y se devuelve el valor. No se puede acceder a los datos directamente. Los datos están ocultos, y eso asegura que no se pueden modificar accidentalmente por funciones externas al objeto.

### **7.3.2 Clases**

Una clase es la descripción de un conjunto de objetos; consta de métodos y datos que resumen características comunes de un conjunto de objetos. Se pueden definir muchos objetos de la misma clase. Dicho de otro modo, una clase es la declaración de un tipo objeto.

Las clases son similares a los tipos de datos y equivalen a modelos o plantillas que describe como se construyen ciertos tipos de objetos. Cada vez que se construye un objeto a partir de una clase, estamos creando lo que se llama una instancia de esa clase. Los objetos no son más que instancias de una clase. Una instancia es una variable de tipo objeto.

#### **7.3.4 Herencia**

Propiedad que permite a los objetos ser contruidos a partir de otros objetos. Dicho de otro modo, la capacidad de un objeto para utilizar las estructuras de datos y los métodos previstos en antepasados o ascendientes. El objetivo final es la reutilizabilidad o reutilización, es decir reutilizar código anteriormente ya desarrollado.

#### **7.3.5 Mensaje**

Un mensaje es una petición de un objeto a otro objeto al que le solicita ejecutar uno de sus métodos. El objeto que envía la petición se denomina emisor y el objeto que recibe la petición se denomina receptor.

Un mensaje consta de tres partes:

- Identidad del receptor.
- El método que ha de ejecutar.
- Información especial necesaria para realizar el método invocado.

#### **7.3.6 Atributos**

Los atributos describen el estado del objeto. Un atributo consta de dos partes: un nombre de atributo y un valor de atributo. Los objetos simples pueden constar de tipos primitivos, tales como enteros, carácter, boolean, reales, o tipos simples definidos por el usuario. Los objetos complejos pueden constar de pilas, conjuntos, listas, array, etc., o incluso estructuras recursivas de alguno o todos los elementos.

### 7.3.7 Métodos

Los métodos describen el comportamiento asociado a un objeto. Representan las acciones que pueden realizarse por un objeto o sobre un objeto. La ejecución de un método puede conducir a cambiar el estado del objeto o dato local del objeto. Cada método tiene un nombre y un cuerpo que realiza la acción o comportamiento asociado con el nombre del método. En un LOO, el cuerpo de un método consta de un bloque de código procedimental que ejecuta la acción requerida. Todos los métodos que alteran o acceden a los datos de un objeto se definen dentro del objeto.

### 7.4 Traductores

Los traductores son un tipo de programas cuya función es convertir el código de un lenguaje en otro. Por ejemplo un **compilador**, que traduce código fuente en código objeto.

Un **traductor** es un programa que recibe como entrada código escrito en un cierto lenguaje y produce como salida código en otro lenguaje. Generalmente el lenguaje de entrada es de más alto nivel que el de salida. Existen distintos tipos de traductores, entre ellos destacan:

- Ensambladores
- Preprocesadores
- Intérpretes
- Compiladores

### 7.5 Ensambladores

Es un tipo de traductor que convierte programas escritos en lenguaje ensamblador en programas escritos en código máquina.

## 7.6 Preprocesadores

Traduce un lenguaje de alto nivel a otro, cuando el primero no puede pasar a lenguaje máquina directamente.

- Procesamiento de macros. Un preprocesador puede permitir a un usuario definir macros, que son abreviaturas de construcciones más grandes.
- Inclusión de archivos. Un preprocesador puede insertar archivos de encabezamiento en el texto del programa.
- Preprocesadores “rationales”. Estos Preprocesadores enriquecen los lenguajes antiguos con recursos más modernos de flujo de control y de estructuras de datos.
- Extensiones a lenguajes. Estos procesadores tratan de crear posibilidades al lenguaje que equivalen a macros incorporales.
- **Intérpretes:** Se trata de traductores-ejecutores ya que con cada instrucción realizan un proceso triple de lectura-traducción-ejecución. Son relativamente lentos, pero muy buenos para la depuración de programas.

Los intérpretes no producen programas objetos, si no que ejecutan directamente los programas fuente. Los intérpretes son populares sobre todo en ambientes de desarrollo de programas en los cuales los programas se ejecutan poco tiempo hasta que aparece un error. Los intérpretes evitan el tiempo adicional del ensamblaje o la compilación. También son muy populares en las computadoras personales.

No obstante, los intérpretes se ejecutan muy lentamente en comparación con el código compilado, ya que debe traducir cada instrucción siempre que esta se ejecuta.

## 7.7 Lenguajes

El concepto de lenguaje no es ajeno para las personas, los estudiosos del lenguaje se han preocupado por imitar las reglas de los lenguajes humanos y

tratar de llevarlas a las máquinas, donde la computadora ha sido un caldo de cultivo muy propicio para obtener excelentes resultados en ese sentido.

Se pueden fragmentar en dos grandes clasificaciones: lenguajes humanos o naturales y lenguajes formales.

El **Lenguaje Natural (LN)** es el medio que utilizamos de manera cotidiana para establecer nuestra comunicación con las demás personas. Este tipo de lenguaje es el que nos permite el designar las cosas actuales y razonar a cerca de ellas, fue desarrollado y organizado a partir de la experiencia humana y puede ser utilizado para analizar situaciones altamente complejas y razonar muy sutilmente.

La riqueza de sus componentes semánticos da a los lenguajes naturales su gran poder expresivo y su valor como una herramienta para razonamiento sutil. Por otro lado la sintaxis de un LN puede ser modelada fácilmente por un lenguaje formal, similar a los utilizados en las matemáticas y la lógica. Otra propiedad de los lenguajes naturales es la poli-semántica, es decir la posibilidad de que una palabra en una oración tenga diversos significados

En un primer resumen, los lenguajes naturales se caracterizan por las siguientes propiedades:

1. Desarrollados por enriquecimiento progresivo antes de cualquier intento de formación de una teoría.
2. La importancia de su carácter expresivo debido grandemente a la riqueza del componente semántico (poli-semántica).
3. Dificultad o imposibilidad de una formalización completa.

El **Lenguaje Formal** es aquel que el hombre ha desarrollado para expresar las situaciones que se dan en específico en cada área del conocimiento científico. Las palabras y oraciones de un lenguaje formal son perfectamente definidas (una palabra mantiene su el mismo significado prescindiendo de su contexto o uso).

Los lenguajes formales son exentos de cualquier componente semántico fuera de sus operadores y relaciones. Los lenguajes formales pueden ser utilizados para modelar una teoría de la mecánica, física, matemática, ingeniería eléctrica, o de otra naturaleza, con la ventaja de que en estos toda ambigüedad es eliminada

En resumen las características de los lenguajes formales son las siguientes:

- Se desarrollan de una teoría preestablecida.
- Componente semántico mínimo.
- Posibilidad de incrementar el componente semántico de acuerdo con la teoría a formalizar.
- La sintaxis produce oraciones no ambiguas.
- La importancia del rol de los números.
- Completa formalización y por esto, el potencial de la construcción computacional.

### 7.8 Puntos para la definición de lenguaje

**Léxico.** Define el conjunto de componentes o palabras que pertenecen al lenguaje.

**Sintaxis.** Define la estructura para formar oraciones o sentencias.

**Semántica.** Evalúa el significado de una oración.

#### ¿Qué es una oración?

La oración es la palabra o conjunto de palabras que tiene sentido completo y autonomía sintáctica, que aunque es imprecisa, trata de reflejar el hecho de que la oración es el fragmento más pequeño del discurso que comunica una idea completa y posee independencia.

## 7.9 Partes de una oración

La gramática tradicional distingue nueve partes de la oración: artículo, sustantivo, pronombre, verbo, adverbio, adjetivo, preposición, interjección y conjunción.

**Artículo** es la parte variable de la oración que se antepone al sustantivo. Se clasifican: definidos e indefinidos.

**Sustantivo** corresponde a una parte que nombra o designa a las personas, animales o cosas.

**Pronombre** son palabras que señalan o sustituyen a otra que normalmente ya se han nombrado.

**Verbo** es una palabra que indica acción, estado de ánimo y acontecimiento de la naturaleza.

**Adverbio** es la palabra no flexiva con función de acompañar y modificar al verbo, también puede modificar al nombre, al adjetivo o a otro nombre.

**Adjetivo** es la parte que acompaña al sustantivo para determinarlo o calificarlo; expresa características o propiedades del sustantivo.

**Preposición** convierte en complemento un sustantivo, un adjetivo, un pronombre, y por otro sirve de introductor de oraciones coordinadas y subordinadas.

**Conjunción** es la que enlaza palabras u oraciones completas, por lo tanto, sirve de nexos o elementos de unión.

El español es una lengua flexiva de tipo fusional, es decir, en las oraciones se usa preferentemente la flexión para indicar las relaciones entre sus elementos. Sin embargo, como la mayoría de las lenguas funcionales, también recurre al uso de preposiciones, palabras abstractas que sirven de nexos y son invariables.

Entre las oraciones encontramos las simples (con un solo núcleo verbal) y las oraciones compuestas (con un núcleo verbal y al menos otro subordinado).

Como idioma flexivo las palabras del español se forman mediante lexemas o raíces a los que se agregan morfemas gramaticales (como el género masculino o femenino y el número singular o plural para los sustantivos y adjetivos, y el modo, tiempo, voz, aspecto y persona y número para el verbo), más todo tipo de afijos que sirven para formar palabras derivadas.

### **7.10 Gramática**

La gramática es un modelo matemático que permite especificar un lenguaje. Mediante el uso de gramáticas, un lenguaje puede describirse a través de un conjunto de pequeñas reglas. En la teoría de compiladores, las gramáticas definen la etapa de sintaxis del lenguaje, operan bajo los conceptos de recursividad.

Las gramáticas describen lenguajes. Los lenguajes naturales como el español, o el inglés, son a menudo descritos por una gramática que agrupa palabras en categorías sintácticas tales como sujetos, predicados, frases preposicionales, etcétera. Expresándolo en forma matemática, una gramática es un dispositivo formal para especificar un lenguaje potencialmente infinito, en una manera finita, puesto que es imposible enumerar todas las cadenas de caracteres en un lenguaje ya sea español, inglés, Al mismo tiempo, una gramática impone una estructura a las sentencias en el lenguaje.

### **7.11 TIPOS DE GRAMÁTICAS JERÁRQUICAS DE CHOMSKY**

Para el estudio de este tema es necesario analizar dos tipos de gramáticas de la clasificación de Chomsky, las regulares y las independientes de contexto, las reglas permitidas y no permitidas.

Tener un conocimiento amplio de las gramáticas y el lenguaje que se emplea en cada una de ellas, es una herramienta más para la realización de los analizadores.

En 1959 Noam Chomsky clasificó las gramáticas en cuatro familias. Las gramáticas no restringidas, sensibles al contexto, independientes del contexto y las gramáticas regulares que se conocen como gramáticas de tipo cero, uno, dos y tres respectivamente. Los lenguajes que resultan de dichas gramáticas también se identifican con lenguajes de tipo cero, uno, dos y tres. A esta jerarquía de lenguaje se le conoce como la jerarquía de Chomsky.

### 7.12 CLASIFICACIÓN DE GRAMÁTICAS

GRAMÁTICA	TIPO	LENGUAJE QUE GENERA
Regular	3	Lenguaje tipo 3 o lenguaje regular.
Independiente de Contexto	2	Lenguaje de tipo 2 o lenguaje independiente de contexto
Sensible al Contexto	1	Lenguaje tipo 1 lenguaje sensible al contexto.
Tipo no Restringido	0	Tipo cero lenguaje no restringido.

Fig. 1.8 Clasificación de gramática.

Figura de Jerarquía de Gramáticas



Fig. 1.9 Jerarquía de gramáticas.

### 7.13 Gramática regular

Es aquella gramática cuyas reglas de reescritura tienen las siguientes restricciones:

1.- El lado izquierdo de cualquier regla de reescritura debe consistir en un solo no Terminal, el lado derecho debe ser un Terminal seguido por un no Terminal, o un solo Terminal o la cadena vacía.

Ejemplo:

$$Z \rightarrow yX$$

$$X \rightarrow y$$

$$X \rightarrow \lambda$$

Las siguientes reglas de reescritura no estarían permitidas en una gramática regular.

Ejemplo:

$$yW \rightarrow X$$

$$X \rightarrow xZy$$

$$YX \rightarrow WvZ$$

En una gramática regular cualquier regla de la forma  $N \rightarrow x$  podría remplazarse con el par de reglas siguientes:

$$N \rightarrow xX$$

$$X \rightarrow \lambda$$

Donde X es un no Terminal que no aparece en ningún otro lugar de la gramática, sin alterar el conjunto de cadenas que podría generar la gramática.

$$N \iff xX \iff x\lambda = x$$

La importancia de las gramáticas regulares reside en que los lenguajes generados por ellas son exactamente aquellos que reconocen los autómatas finitos.

**Nota:**

→ Se interpreta como "**puede ser**", "se compone de", "es sustituida por".

\ se interpreta como "**o**" ejemplo:  $E \rightarrow A$  y  $E \rightarrow B$  se unen en  $E \rightarrow A|B$ .

↪ interpreta como "**derivar**", "**produce**" o "**genera**".

**7.14 Gramáticas independientes del contexto**

A diferencia de las gramáticas regulares, estas gramáticas no tienen restricciones con respecto a la forma del lado derecho de sus reglas de reescritura aunque aún se requiere que el lado izquierdo de cada regla sea un no Terminal. La siguiente es una gramática independiente del contexto, pero no es regular.

**S** → **zMNz**

**M** → **aMa**

**M** → **z**

**N** → **bNb**

**N** → **z**

El término independiente del contexto refleja que, como el lado izquierdo de cada regla de reescritura únicamente puede contener un solo no Terminal la regla puede aplicarse sin importar el contexto donde se encuentre dicho no Terminal. Al igual que las gramáticas regulares, las independientes del contexto (G.I.C.) generan cadenas por medio de derivaciones. En el caso de las G.I.C. pueden surgir dudas con respecto a cuál será el no Terminal que deberá reemplazarse en un paso específico de la derivación.

**7.15 Gramáticas sensibles al contexto**

Las gramáticas sensibles al contexto o de tipo 1, son las que generan los lenguajes sensibles al contexto. Los lenguajes sensibles al contexto son aquellos que pueden ser reconocidos por las Autómatas Linealmente Acotados ALA.

---

En forma general toda gramática se define mediante una cuádrupla  $G=(N,T, P,S)$ , siendo.

-**N** es un conjunto finito de símbolos no terminales

-**T** es un conjunto finito de símbolos terminales  $N \cap T = \emptyset$

-**P** es un conjunto finito de reglas de producción

-**S** Símbolo distinguido o Axioma  $S \in (N \cap T)$

### 7.16 Gramáticas no restringidas

- Permiten producciones que eliminan símbolos.
- Permite a un símbolo expandirse y contraerse.
- Máquina de Turing.

### 7.17 Gramática del español

La **Gramática** es el estudio de las reglas y principios que regulan el uso de las lenguas y la organización de las palabras dentro de una oración. También se denomina así al conjunto de reglas y principios que gobiernan el uso de un lenguaje muy determinado; así, cada lenguaje tiene su propia gramática.

### 7.18 Términos Gramaticales:

#### 7.18.1 Adjetivo

Es una parte de la oración que acompaña al sustantivo o nombre para determinarlo o calificarlo; expresa características o propiedades atribuidas a un sustantivo, ya sean concretas (el libro **verde**, el libro **grande**), ya sean abstractas (el libro **difícil**). Estos adjetivos acompañan al sustantivo libro y cumplen la función de especificar o resaltar alguna de sus características y se dice que lo determinan, pues, al añadir un adjetivo ya no se habla de cualquier libro, sino precisamente de un libro verde, o de uno grande.

### 7.18.2 Artículo

Sirve para precisar la referencia de un sustantivo, transformándolo de desconocido y abstracto ("libro") a conocido y concreto ("el libro"). Dicho de una forma más clara, es un implemento o indicador que sitúa lo que se dice en el campo en que se dice, común al que lo dice y al que lo oye, o sea, en el mutuo acuerdo de los interlocutores.

En el castellano hay unos 11 artículos diferentes en singular y plural. Se clasifican así:

Artículos gramaticales en castellanos variables.

	Masculino singular	Femenino singular	Masculino Plural	Femenino Plural
Determinados	El	La	Los	Las
Indeterminados	Un	Una	Unos	Unas

Tabla 1.3 Artículos variables.

Artículos gramaticales en castellanos invariables.

Neutro	Lo	
Contractos	Al (a +el)	Del (de + el)

Tabla 1.4 Artículos invariables.

### 7.18.3 Adverbio

El **adverbio** es la clase de palabra que actúa como núcleo del sintagma adverbial. Modifica al verbo, a un adjetivo o a otro adverbio.

Las funciones sintácticas del adverbio son, aparte de la de núcleo de sintagma adverbial, las de complemento circunstancial del verbo, la de cuantificador, grado o complemento del adjetivo ("muy bueno", "recién hecho") y las de cuantificador de otro adverbio ("bastante lejos"). Algunos adverbios pueden funcionar como predicados dirigidos hacia un sujeto y junto a una cópula verbal ("está divinamente").

### **Adverbio**

Se conoce como adverbio a aquella parte de la oración, generalmente representada en una palabra y que modifica el significado del verbo de la oración en cuestión o bien de las otras palabras como puede ser un adjetivo o asimismo otro adverbio.

Las funciones de los adverbios son básicamente las de complemento circunstancial del verbo de una oración, las de cuantificador, grado o complemento en el caso de los adjetivos y las de cuantificador si se trata de otro adverbio.

### **Ejemplo:**

Este paraguas **no sirve**. (*El **no** es el adverbio y el **sirve** es el verbo*).

### **Adverbios y sus categorías**

Tradicionalmente los adverbios se han dividido en las siguientes categorías:

Adverbios de ubicación en tiempo y espacio:

**Adverbios de lugar:** aquí, allí, ahí, allá, acá, arriba, abajo, cerca, lejos, delante, detrás, encima, debajo, enfrente, atrás, alrededor, etc. y así son algunos ejemplos

**Adverbios de tiempo [absoluto]:** pronto, tarde, temprano, todavía, aún, ya, ayer, hoy, mañana, siempre, nunca, jamás, próximamente, prontamente, anoche, enseguida, ahora, mientras.

**Adverbio de orden [tiempo relativo]:** antes, después, posteriormente, primeramente, primero, respectivamente

Adverbios de modo y cantidad:

**Adverbios de modo:** bien, mal, regular, despacio, deprisa, así, tal, como, aprisa, adrede, peor, mejor, fielmente, estupendamente, fácilmente - todas las que se formen con las terminaciones "mente".

**Adverbios de cantidad o grado:** muy, poco, muy poco, cada vez más, mucho, bastante, más, menos, algo, demasiado, casi, sólo, solamente, tan, tanto, todo, nada, aproximadamente.

Adverbios que cumplen la función de epistémicos:

**Adverbios de afirmación:** sí, también, cierto, ciertamente, efectivamente, claro, exacto, verdaderamente.

**Adverbios de negación:** no, jamás, nunca, tampoco, negativamente...

**Adverbios de duda o dubitativos:** quizá(s), acaso, probablemente, posiblemente, seguramente, tal vez, sin duda.

#### 7.18.4 Verbo

El Verbo: El verbo es el núcleo del predicado y es la acción del sujeto.

- Este consta de tres modos: Indicativo, subjuntivo e imperativo
- Tiene tres tiempos: presente, pasado y futuro.
- Número: singular y plural
- Tres personas: 1ª yo/nosotros 2ª tú/vosotros 3ª él/ellos.
- Consta de tres conjugaciones: 1ª -ar, 2ª -er, 3ª -ir.

Ejemplo: el verbo, **caminar**.

## **8. HIPÓTESIS**

En este proyecto se desarrolla un sistema que traducirá oraciones simples en español para poder crear la gramática en español al tsotsil, se usara una gramática de tipo LR que nos permitirá traducir oraciones simples en tiempo presente de español a tsotsil y oraciones simples en tiempo pasado de español a tsotsil.

## 9. ANÁLISIS DEL SISTEMA

### 9.1 GRAMÁTICA TSOTSIL

Para fines de este proyecto se identificaron 9, en este orden:

1. Sustantivo
2. Adjetivo Calificativo
3. Pronombre Personal
4. Marca Posesiva
5. Artículo Determinante
6. Marca de Sujeto
7. Preposición
8. Marca Aspectual de Tiempo
9. Verbo

### 9.2 Sustantivo

Es aquel que significa alguna sustancia corpórea o incorpórea como: *na* (casa), *te'* (árbol), *kerem* (niño), *totil* (papá). Subsiste por sí mismo en la oración. Sin necesidad de que se le junte otra palabra que le califique. Los sustantivos se clasifican en dos tipos:

- Sustantivo común

Ejemplo: *lum* (pueblo), *totil* (señor), *uk'um* (río).

- Sustantivo propio

Ejemplo: *jxun* (Juan), *jxunka'* (Juana).

### Adjetivo calificativo

Es el que se junta al sustantivo para denotar su calidad, como: *lek* (bueno), *chopol* (feo), *sak* (blanco), *ik'* (negro). El adjetivo no puede estar en la oración sin sustantivo expreso o suplido.

Está expreso cuando decimos: *lek li vinike* (hombre bueno) y suplido cuando decimos: *ta sbulan ch'ulna li leke* (el bueno visita la iglesia) porque se sule el sustantivo hombre. En estos casos se dice que los adjetivos están sustantivados, o se usan como sustantivos. Los adjetivos son palabras que califican a los sustantivos, describen características del sujeto u objeto directo. Por lo general en tsotsil, tradicionalmente se escribe la palabra adjetival antes de la palabra nominal. Los adjetivos no se pluralizan en su propia palabra, es decir, se pluraliza a la palabra nominal sustantiva, por ejemplo:

*Jujp'em vinik* → hombre gordo.

*Jujp'em viniketik* → hombres gordos.

*Jujp'em* significa gordo *vinik* hombre, entonces al pluralizar el sustantivo en *viniketik*, ya estoy pluralizando el adjetivo, quedando:

*jujp'em viniketik* → hombres gordos.

### 9.3 Marca posesiva

Son los que denotan posesión, o pertenencia de alguna cosa, como: *jna* (mi casa), *jtot* (mi papá).

Las Marcas Posesivas son aquellos prefijos que le anteceden exclusivamente a los sustantivos o adjetivos calificativos.

La lengua tsotsil tiene sus propias características que es mediante el uso de prefijos que trata de posesionar todo aquello que se enuncia.

Encontramos dos grupos:

1. Los que acompañan a los sustantivos que inician con vocal (**k, av, y**).
2. Los que van antes de la raíz del sustantivo que comienzan con consonantes (**j, a, s o ta, cha, ts**).

#### 9.4 Artículo determinante

En el idioma español es una parte de la oración que sirve para distinguir los géneros de los nombres. Para los masculinos sirven *el* en singular, y *los* en plural: para los femeninos *la*, en singular, y *las* en plural.

*En la lengua tsotsil existe un sólo artículo que marca para todo el género: masculino y femenino, que es la palabra "li"; además, no modifica su estructura si es singular o plural.*

Y así decimos: **li** *vinike* (**el** hombre), y **li** *viniketike* (**los** hombres): **li** *antse* (**la** mujer), y **li** *antsetike* (**las** mujeres).

#### 9.5 Marca de sujeto

Son los que denotan posesión, o pertenencia de alguna acción, como: **cha** *ve'* (comes), **cha** *bit*(brincas), **jmantik** (compramos).

Las Marcas de Sujeto son aquellos prefijos o sufijos que exclusivamente acompañan a los verbos y permiten la conjugación de los mismos.

La marca de sujeto va pegado o separado al verbo y va a depender si el verbo empieza con vocal o con consonante si el verbo empieza con consonante siempre se usa un espacio o un (-).

#### 9.6 Preposición

En idioma español es una palabra llamada así, porque se pone antes de otras partes de la oración.

El oficio de la preposición por sí sola es indicar en general alguna circunstancia que no se determina sino por la palabra que se le sigue; pero junta ya con ella, denota la diferente relación o respeto que tienen unas cosas con otras, como: *á, con, de, en, por, para*.

En la lengua tsotsil solo existe una preposición **ta** que significa *a, de, en, en el, en el, y a*. Por ejemplo:

*ta xave' ta mexa* (comes en la mesa).

*ta xanab ta yaxaltik* (camina en el campo).

### 9.7 Marca aspectual de tiempo

En el idioma español hay solo tres tiempos, que son: presente, pasado, y venidero; esos mismos conoce la Gramática en los verbos, y los llama: presente, pretérito, y futuro.

En la lengua tsotsil se contemplan actualmente la representación de solo dos tiempos los cuales son:

La marca de tiempo **ta** (*presente*) que denota lo que es, se hace, ó sucede actualmente, como:

*ta sts'i'baj* (escribe), *ta sve'* (come).

La marca de tiempo **la** (*pasado*) demuestra que alguna cosa fue, se cumplió, ó sucedió, como:

*la kuch* (cargaste), *la sxut'* (pellizcó).

La marca de tiempo futuro no existe como tal sino que se interpreta que la acción se realizará pronto, este tiempo se forma con el morfema **ta to** equivalente a todavía o aún, como:

**ta to jman** (compraré).

**ta to xavil** (verás).

## Verbos

En tsotsil todos los verbos se presentan de forma infinitiva como “manel” que significa comprar.

El sufijo **el** indica que el verbo se encuentra en su forma infinitiva y por lo tanto no se encuentra conjugado. Al conjugarse pierden ese sufijo. Los verbos se clasifican en dos tipos: verbos transitivos y verbos intransitivos:

**9.8 Verbo transitivo:** Es aquel que transfiere la acción a una persona gramatical (objeto, persona, cosa, etc.). Para conjugar a los verbos intransitivos tenemos que saber si el verbo empieza con vocal o con consonante por ejemplo:

manel = comprar. Este verbo comienza con consonante por lo tanto le corresponde la siguiente conjugación.

jman = yo compro	}	singular
aman = tu compras		
sman = él compra		

jmantik = nosotros compramos	}	plural
amanik = ustedes compran		
smanik = ellos compran		

**9.9 Verbo intransitivo:** Son aquellas acciones que no son transferibles a personas gramaticales por lo que es una sola persona quien realiza la acción y recae en si mismo.

A diferencia de los verbos transitivos aquí ya no se emplea la “j” y “k” para las primeras personas del singular sino se emplea el sufijo “on” (YO), por ejemplo:

ve'**el** = comer

Conjugación:

<b>Chi</b> ve' = yo como	}	singular
<b>Cha</b> ve' = tu comes		
<b>Ch</b> ve' = él come		

<b>Chi</b> jve' <b>otik</b> = nosotros comemos	}	plural
<b>Cha</b> ve' <b>ik</b> = ustedes comen		
<b>ch</b> ve' <b>ik</b> =ellos comen		

### 9.10 Clítico

**Li** tiene la categoría de artículo determinante, antes de la palabra nominal. Tiene la categoría de palabra independiente sin embargo al término de la oración se sufija el sonido **e**, mismo que indica el final de la oración o frase y a este se le denomina clítico.

Por ejemplo:

kerem (niño)

**li** kereme (el niño)

### 9.11 Verbos multicategoría

Existen verbos en español que tienen más de una categoría en la lengua Tsotsil por ejemplo:

Comer, alto, cortar entre otros.

En español comer es un término general y en tsotsil no son más específicos, dependiendo de qué es lo que se está comiendo, por ejemplo:

El señor **come** carne → **Sti'** v'ek'et li vinike.

El señor **come** tortilla → **tsve'** vaj li vinike.

El señor **come** pera → **slo'** pelax li vinike.

El señor **corta** leña → **stuch'** si' li vinike.

El señor **corta** repollo → **tset** itaj li vinike.

### 9.12 Sufijo del sustantivo

Un sustantivos en su forma normal se presentan con un sufijo *il* el cual indica que el sustantivo no se posee por ejemplo:

*totil* (*papa*) al poseer el sustantivo este sufijo desaparece quedando *jtot* (*mi papá*).

### 9.13 Gramática de oraciones simples en tiempo presente de español a tsotsil

Estructura de una oración simple en español.

**Sujeto + predicado**

---

Estructura de una oración simple en tsotsil.

### Predicado + sujeto

En el marcador de tiempo en presente solo se pone antes del verbo **chi** para primera persona, **cha** para segunda persona **ch** para tercera persona y en pasado una **li** para primera persona y **la** para segunda persona y para la tercera persona no se agrega nada, pero no se aplica en todas solo cuando se trata de personas físicas y directas por ejemplo:

Corriste en el campo.

**La** anilaj ta yaxaltik

Como se puede ver la estructura de la oración en tsotsil es lo contrario la del español, por lo tanto la gramática quedaría de la siguiente manera.

#### 9.14 Gramática tsotsil en tiempo presente.

<ORACIÓN> → <PREDICADO> <SUJETO>

<PREDICADO> → **chi** | **cha** | **ch** <VERBOCPRE><COMPLEMENTO> | □

<VERBOC> → <verboc>

<COMPLEMENTO> **ta** <ADJ> <PGRAMATICAL> | <ADJ> <PGRAMATICAL> | □

<ADJ> → <adj> | □

<PGRAMATICAL> → <MP><SUST> | <SUST>

<MP> → **j** | **a** | **s** | **k** | **av** | **y** | □

<SUST> → <sust>

<SUJETO> **li** <ADJ><PGRAMATICAL> | □

**VERBOCPRE:** Verbo conjugado en presente (lexicon)

**ta:** Preposición.

**chi:** Marca aspectual de tiempo primera persona

**cha:** Marca aspectual de tiempo segunda persona.

**Ch:** Marca aspectual de tiempo tercera persona.

**ADJ:** Adjetivo (lexicon).

**PGRAMATICAL:** Persona gramatical.

**MP:** Marca posesiva.

**SUST:** Sustantivo (lexicon).

**LI:** Artículo determinante.

□: Vacío.

### 9.15 Gramática tsotsil en tiempo pasado.

<ORACIÓN> → <PREDICADO> <SUJETO>

<PREDICADO> → **li | la** <VERBOCPAS><COMPLEMENTO> | □

<VERBOC> → <verboc>

<COMPLEMENTO> **ta** <ADJ> <PGRAMATICAL> | <ADJ> <PGRAMATICAL> | □

<ADJ> → <adj> | □

<PGRAMATICAL> → <MP><SUST> | <SUST>

<MP> → **j | a | s | k | av | y** | □

<SUST> → <sust>

<SUJETO> **li** <ADJ><PGRAMATICAL> | □

**VERBOCPAS:** Verbo conjugado en pasado (lexicon)

**li:** Marca aspectual de tiempo en pasado para primera persona.

**la:** Marca aspectual de tiempo en pasado para segunda persona.

**ta:** Preposición.

**ADJ:** Adjetivo (lexicon).

**PGRAMATICAL:** Persona gramatical.

**MP:** Marca posesiva.

**SUST:** Sustantivo (lexicon).

**LI:** Artículo determinante.

□: Vacío.

### 9.16 Gramática español adaptada al tsotsil con oraciones simples en presente y pasado.

<ORACIÓN> → <SUJETO> <PREDICADO>

<SUJETO> **li** <ADJ><PGRAMATICAL> | □

<PGRAMATICAL> → <MP><SUST> | <SUST>

<MP> → **j | a | s | k | av | y** | □

<SUST> → <sust>

<ADJ> → <adj> | □

<PREDICADO> → **li | la** | □ <VERBOCPAS><COMPLEMENTO> | **chi | cha | ch**

<VERBOCPRE><COMPLEMENTO> | □

<VERBOC> → <verboc>

---

<COMPLEMENTO> **ta** <ADJ> <PGRAMATICAL> | <ADJ> <PGRAMATICAL> | □

**VERBOCPAS:** Verbo conjugado en pasado (lexicon).

**VERBOCPRE:** Verbo conjugado en presente (lexicon).

**chi:** Marca aspectual de tiempo en presente para primera persona.

**cha:** Marca aspectual de tiempo en presente para segunda persona.

**ch:** Marca aspectual de tiempo en presente para tercera persona.

**li:** Marca aspectual de tiempo en pasado para primera persona.

**la:** Marca aspectual de tiempo en pasado para segunda persona.

**ta:** Preposición.

**ADJ:** Adjetivo (lexicon).

**PGRAMATICAL:** Persona gramatical.

**MP:** Marca posesiva.

**SUST:** Sustantivo (lexicon).

**LI:** Artículo determinante.

□: Vacío.

Esta gramática denominada del español adaptada al tsotsil, contiene la estructura de las oraciones en español, pero con los elementos que contempla la gramática del tsotsil, es decir, contiene la estructura general de las oraciones en español es <ORACION> → <SUJETO> <PREDICADO>, considerando que la cadena de entrada está escrita en español esta debe ser la estructura que tendrán, pero para traducirla tendremos que adaptarla a los elementos de la oración en tsotsil, que contiene una estructura general inversa (<ORACION> → <PREDICADO> <SUJETO> ); en la producción (<SUJETO> → **ta** <PGRAMATICAL> <ADJ> |  $\epsilon$  ),

nótese que después de la persona gramatical (<PGRAMATICAL>) está el adjetivo (<ADJ>).

### 9.17 Ejemplos de ORACIONES SIMPLES en Tsotsil

Los nombres o sustantivos funcionan también “verbos de estado” o “predicados no verbales”, agregándoles el sufijo **un** para la primera persona, **ot** para la segunda persona y **no agregar nada** para la tercera persona y la **J** que va en el prefijo del verbo sirve cuando están poseídos por una persona, por ejemplo :

Estudiante = chanun.

Soy estudiante

vu'un **jchanunun**

Escritor = ts'ibajom

- Yo soy un escritor =vu'un **jtsi'bajomun** (la **j** ya se sabe que es cuando están poseídos los verbos)
- Tu eres escritor = vo'ot **Jtsi'bajomot**
- El (ella) es escritor (a) = ja' **jtsi'bajom**
- Nosotros (as) somos escritores (as) =vo'otik **jtsi'bajomotik** (incluyendo el interlocutor).
- Nosotros (as) somos escritores (as) =vu'unikutik **jtsi'bajomunkutik** (excluyendo al interlocutor).
- Ustedes son escritores (as) = vo'oxuk **jtsi'bajomoxuk**.

Ellos (as) son escritores (as) =ja'ik **Jtsibajometik**.

### Ejemplos de oraciones en tiempo presente.

El caballo corre en el campo.

El (**artículo**) caballo (**sujeto**) corre (**verbo**) en el campo (**complemento**).

Li ka'e ch anilaj ta yaxaltik

Li (**artículo** y siempre se pone **e** al final del sujeto) ka'e (**sujeto** "caballo= ka'") ch anilaj (**verbo**) ta yaxaltik (**complemento**).

Cuando entra un adverbio es cuando altera el orden de la oración como se muestra en el siguiente ejemplo:

El caballo corre "rápido" en el campo.

El (**artículo**) caballo (**sujeto**) corre (**verbo**) rápido (**adverbio**) en el campo (**complemento**).

Li ka'e anil ch anilaj ta yaxaltik

Li (**artículo**) ka'e (**sujeto** "caballo= ka'") anil (**adverbio**) ch anilaj (**verbo**) ta yaxaltik (**complemento**).

### **Ejemplo de oraciones en tiempo pasado.**

Él estuvo trabajando = Le'e abtej to'ox.

Él (**sujeto**) estuvo (**ser o estar en pasado**) trabajando (**verbo**).

Le'e (**sujeto**) abtej (**verbo**) to'ox (**ser o estar en pasado**).

Yo estuve comiendo = vu'un li ve' to'ox.

Yo (**sujeto**) estuve (**ser o estar**) comiendo (**verbo**).

Vu'un (**sujeto**) li ve' (**verbo**) to'ox (**ser o estar**).

Tenía tres caballos = Oy xa ox oxkot jka'.

Tenía (**verbo**) tres (**adverbio cantidad**) caballos (**objeto**).

Oy xa ox (**tenía**) oxkot (**adverbio cantidad**) jka' (**objeto**).

**Oy** no es un verbo; indica la existencia o pertenencia de algo.

**Xa ox** lo vuelve pasado porque el **oy** se puede manejar en cualquier tiempo ya se presente o pasado.

Había maíz = oy to ox ixim.

El alfabeto está compuesto de veinticinco grafías, de las cuales veinticinco son consonantes y cinco vocales. El alfabeto Tsotsil utilizado es el siguiente:

**[ a b ch ch' e i j k k' l m n o p p' r s t t' ts ts' u v x y ]**

Las grafías consonánticas; b, ch, j, k, l, m, n, p, r, s, t, y, más las 5 vocales, representan sonidos equivalentes a los de la lengua Español.

La grafía **k** sustituye las consonantes **c** y **qu** del español y se usa dentro de cualquier vocal.

La grafía **x** es equivalente al sonido de la **x** en una gran cantidad de las palabras.

La **ts** representa un sonido combinado de los expresados por los de la grafía **t** y **s**, pero recientemente en los últimos decenios un importante número de escritores indígenas Tsotsiles ha acordado emplear **ts**.

El apóstrofo o sonido “glotal”, a menudo llamado “el glotal” como se discrimina en la lingüística (‘) se produce mediante un cierre de las cuerdas vocales, lo que da como resultado un sonido explosivo para los casos de la grafías **ch'**, **k'**, **t'**, **ts'**, representando los sonidos consonánticos vocales y que son en efecto explosivos.

### **Ejemplos:**

En el siguiente ejemplo se podrá ver que el glotal o apostrofe influye mucho, se ocupan las mismas letras pero con poner el glotal cambia el significado.

T'ot' = caracol

Tot = papá

En la escritura de algunas palabras se utiliza el (-), pero sólo en aquellas donde una consonante se pronuncia independiente de la vocal que la sigue, como ejemplo en:

j-abtel = trabajador.

x-avet = gritando.

Español	Tsotsil
Yo	Vu'un
Tu	Vo'ot
El	Le'e
Nosotros	vo'otik
Ustedes	vo'oxuk
Ellos	Le'ik

Tabla 1.5 Pronombres personales.

### Presente simple

Estructura:

Sujeto + verbo + complemento.

#### Ejemplos:

Estoy caminando en el parque. (**Estoy** es el sujeto, **caminando** verbo **en el parque** complemento).

En Tsotsil equivale a:

Ta xi xanab ta ch'ivit. (**vu'un chi** es el sujeto, **xanab** verbo **ta chi'vit** es el complemento).

Cuando se ocupa algún adverbio en español, se altera el orden de la oración en la lengua Tsotsil al hacer la traducción, como se muestra en el siguiente ejemplo:

**Eres poco hablador.** (**Eres** es el sujeto, **poco** adverbio, **hablador** verbo)

**Jutuk xa na' xa k'opoj.** (**Jutuk** es el adverbio, **xa na'** el sujeto, **xa k'opoj** verbo)

Yo camino todo los días (**Yo** es el sujeto, **camino** verbo, **todos los días** el complemento).

**Ta xi xanab skotol k'ak'al** (*Ta el sujeto, xi xanab verbo, skotol k'ak'al complemento*).

Se maneja con el nombre de **préstamo** a las palabras que aún no están adaptadas totalmente, porque conservan los sonidos originales de la lengua que proviene.

**Kilo**

**Televisión**

**Chicle**

Ta jk'el **televisión** = mirando la **televisión**

Ta jman **chikle** = comprando **chicle**.

Jun **kilo** askal = un **kilo** de azúcar.

### 9.18 DIAGRAMA DE CASOS DE USO

El usuario interactúa con el sistema, en el inicio cuando se ejecuta el sistema traductor, tiene las opciones de introducir el texto a traducir en español o ir a la barra de herramientas para abrir archivos de texto o cambiar fuente de letras también usar el diccionario entre otras cosas, en traducir tiene las opciones de imprimir la oración traducida o guardar en un archivo \*.txt o en \*.pdf, al terminar el usuario tiene la opción de salir o empezar con otra traducción, como se muestra en la figura 1.11

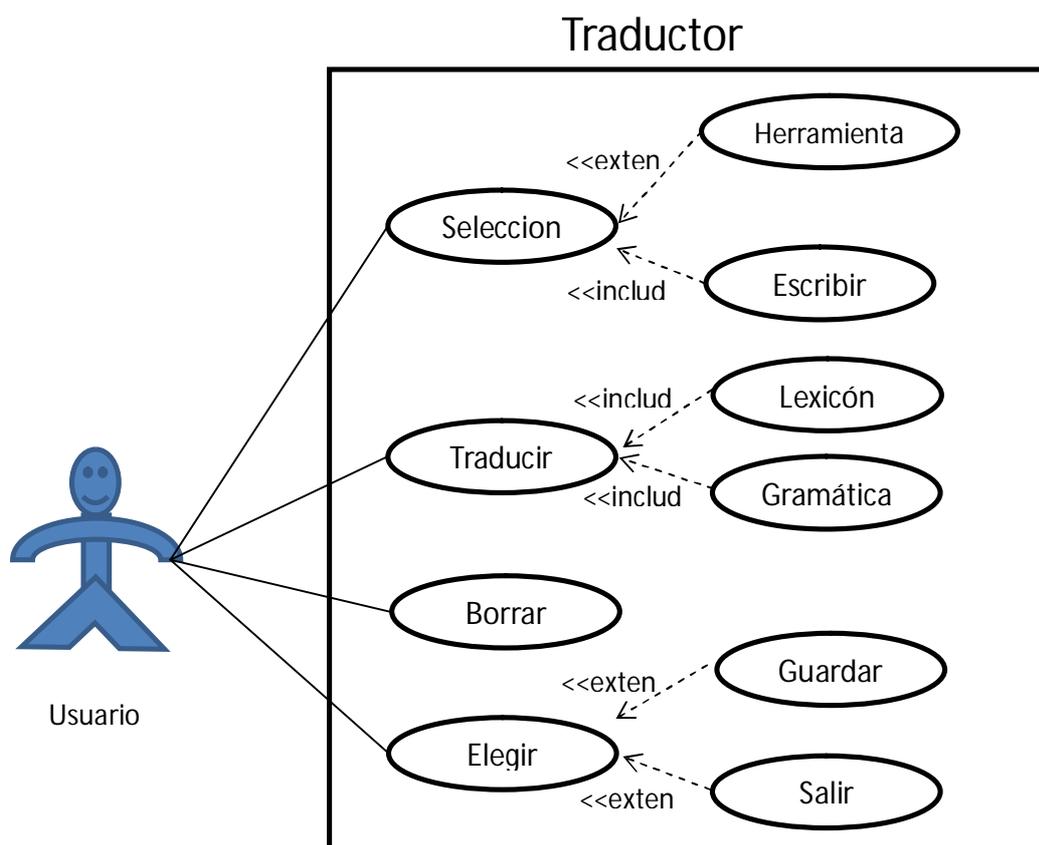


Fig. 1.10 Diagrama de casos de uso.

### 9.18.1 Descripción de los casos de uso

Nombre del caso de uso	Selección	
Actor	Usuario	
Descripción	El usuario puede seleccionar cualquiera de las herramientas disponibles en la barra de tareas de herramienta del traductor.	
Acción	Actor	Sistema
	1.-Elige las opciones que existe en la barra de herramientas o escribe una oración simple en tiempo presente o pasado.	1.- muestra las herramientas disponibles en el sistema o las opciones para escribir la oración.

Nombre del caso de uso	Traducir	
Actor	Usuario	
Descripción	El usuario puede activar el botón traducir.	
Acción	Actor	Sistema
	1.-Activa el botón traducir.	1.- Divide la cadena en subcadenas. 2.- etiqueta las subcadenas con la etiqueta correspondiente al lexicón. 3.- se hace la traducción al tsotsil en presente o pasado. 4.-Ordena la oración con la gramática tsotsil. 5.- ordena la oración. 6. imprime en pantalla.

Nombre del caso de uso	Borrar	
Actor	Usuario	
Descripción	El usuario puede borrar la cadena introducida.	
Acción	Actor	Sistema
	1.-Activa el botón borrar.	1.- Limpia el cuadro de texto en tsotsil y español.

Nombre del caso de uso	Elegir	
Actor	Usuario	
Descripción	El usuario tiene las opciones de elegir que formato quiere guardar la oración, imprimir o simplemente borrar de nuevo.	
Acción	Actor	Sistema
	1.- Selecciona la opción para guardar la oración traducida, también puede mandar a imprimir la oración o borrar la oración.	1.-Muestra las opciones para guardar la oración y la opción de mandar a imprimir o simplemente borrar.

## 9.19. DISEÑO DEL SISTEMA

En este apartado como se muestra en la fig. 1.11, se puede observar el diseño detalladamente del sistema traductor del español a tsotsil.

### 9.19.1 Diagrama de clases:

Se describe la estructura del sistema de traducción a través de un diagrama de clases, mostrando los campos y métodos de cada clase.

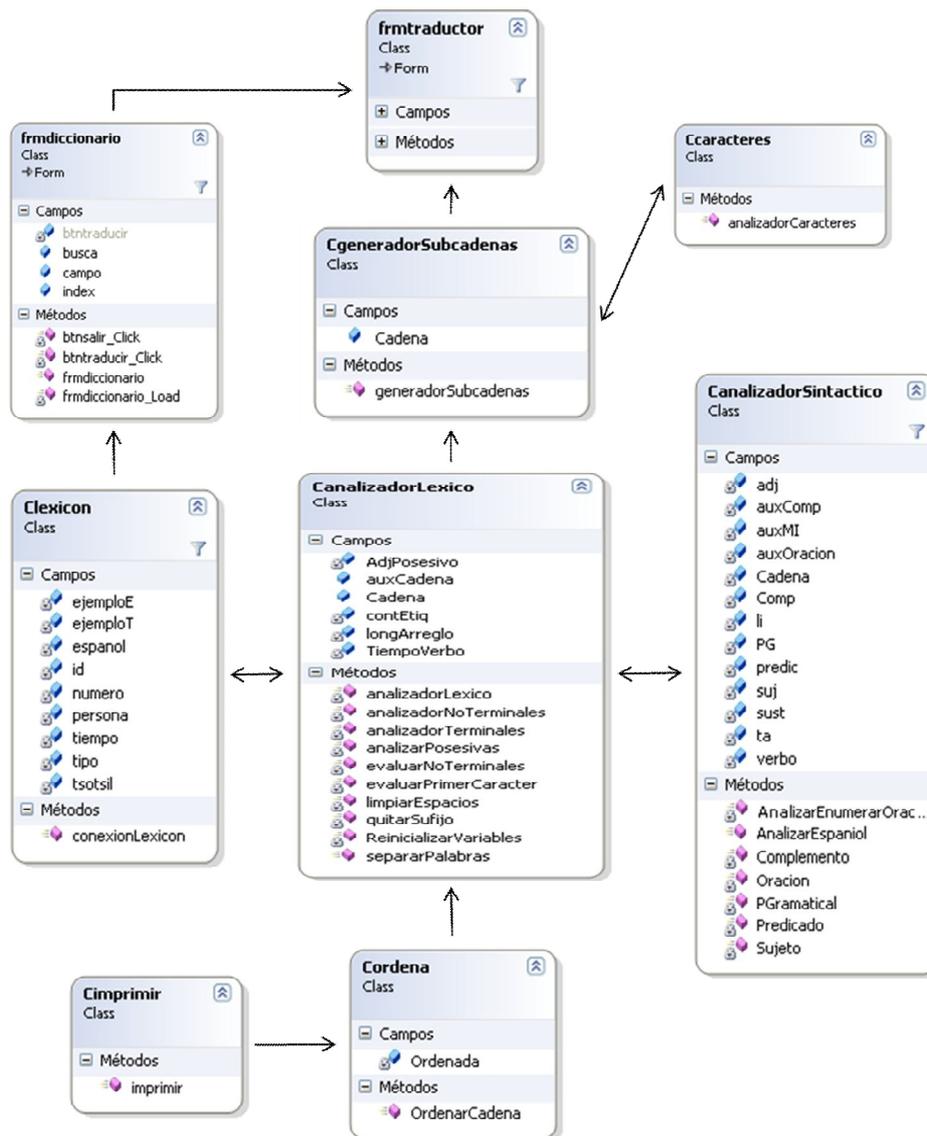


Fig. 1.11 Diagrama de clases.

### **DIAGRAMA DE CLASE DEL TRADUCTOR TSOTSIL.**

La clase `fmrtraductor` es la interfaz del traductor encargada de mostrar las acciones que se realizan en el traductor.

La clase `CgeneradorSubcadenas` se encarga de separar una cadena en subcadenas, toma como una cadena hasta donde encuentre un delimitador.

La clase `Ccaracteres` se encarga de verificar si entra algún carácter raro o inesperado para marcar como error en la oración.

La clase `CanalizadorLexico` se encarga de analizar si las subcadenas son terminales o no terminales; también analiza aquellas palabras desconocidas, es decir, aquellas palabras que no se encuentran en el lexicón; además hace un análisis de los verbos y sustantivos que se le tiene que quitar el sufijo al formar una oración en tsotsil.

La clase `CanalizadorSintactico` (gramática del tsotsil) se encarga de establecer el orden de las subcadenas para que la oración en tsotsil esté formada correctamente.

La clase `Clexicón` se encarga de establecer la conexión con la base de datos (lexicón), que es donde se encuentran los verbos, sustantivos, etc., con su equivalente en tsotsil.

La clase `Cordena` es quien se encarga de formar la oración correcta en tsotsil después de que la gramática haya establecido el orden en que ha de formarse dicha oración en tsotsil.

La clase `Imprimir` se encarga de mandar a pantalla la oración traducida correctamente en tsotsil.

La clase `fmrDiccionario` es la interfaz del diccionario encargado de mostrar la traducción de las palabras.

### 9.19.2 Estructura del lexicón

En esta sección visualizaremos como está constituida nuestra base de datos como por ejemplo sus tablas, campos y registros.

En el diseño de este sistema traductor se usó Access para la base de datos y se le denomina **lexicón** dentro del lexicón podemos encontrar verbos, sustantivos, adjetivos y adverbios, todos ellos escritos en español y tsotsil como se muestra en la figura 1.12

Tsotsil	Español	Etiqueta	Tiempo	Persona	Número	EjemploT	EjemploE
Vayel	Dormir	VT	0	0		Li ka'e	El caballo
Chi vay	Duermo	VT	PRE	1	S		
Ch vay	Duerme	VT	PRE	2	S		
Vay	Dormí	VT	PAS	1	S		
Vay	Durmió	VT	PAS	2	S		

Fig. 1.12 Estructura del lexicón.

## 10. DESARROLLO DEL SISTEMA

En la figura 1.13 se muestra el modelo conceptual y cada uno de sus componentes que son:

Entrada de archivo de texto en español, que puede ser una oración simple en tiempo presente o pasado.

Generador de subcadenas, donde se almacenan las palabras introducidas para después validarlos.

El siguiente paso es pasar al analizador léxico, que es la parte donde se evalúan las componentes o las palabras si pertenecen al lenguaje.

Después la oración introducida pasará a un analizador sintáctico, que es el encargado de ver las estructuras de las oraciones para poder pasarlos a tsotsil en forma correcta interactuando con el lexicón que es donde se encuentra almacenado todos los datos para después evaluarlos con la gramática de tsotsil que se encuentra en el proceso de refinamiento después imprimirlo.

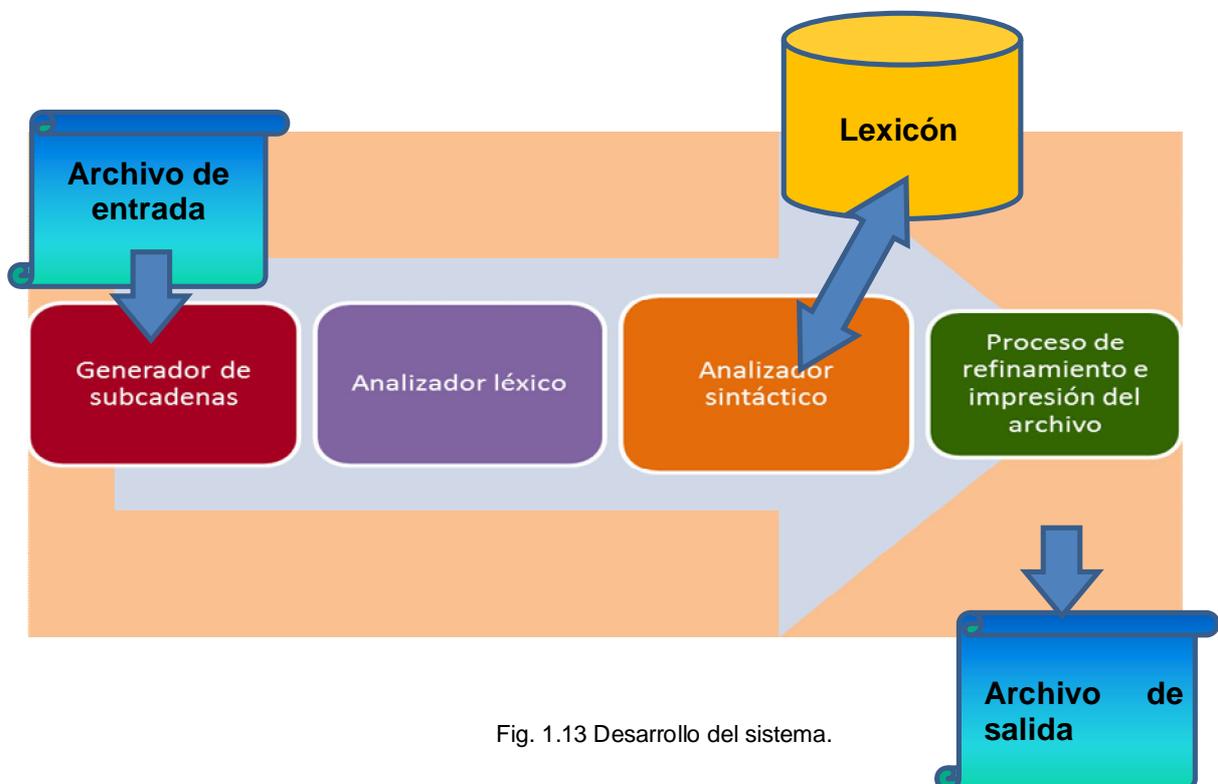


Fig. 1.13 Desarrollo del sistema.

El funcionamiento del sistema traductor de español al tsotsil, es que mediante la entrada de un archivo o una cadena de caracteres, pase al generador de sub cadenas, el cual es el encargado de verificar que las oraciones o texto vaya bien escritos conforme a la gramática en español, después pasa al analizador léxico que trabaja conjuntamente con el analizador sintáctico, los cuales funcionan para verificar si existe algún delimitador para poder traducir el texto a la lengua tsotsil, traducirá con las palabras adecuadas que está almacenado en el lexicón, que es la base de datos donde se encuentran todas las palabras en español y tsotsil para una traducción adecuada, después pasamos a la impresión del texto ya en tsotsil con el orden adecuado, y por ultimo pasa a la impresión del archivo ya sea \*.txt, \*.pdf con el archivo o texto traducido correctamente en Tsotsil.

### 10.1 Diagrama de bloques (Modelo Conceptual del Sistema):

Se representa gráficamente el funcionamiento interno del sistema traductor, aquí se definen todos los procesos que lleva a cabo el sistema, sus entradas y salidas.

#### 10.1.1 DIAGRAMA A BLOQUES

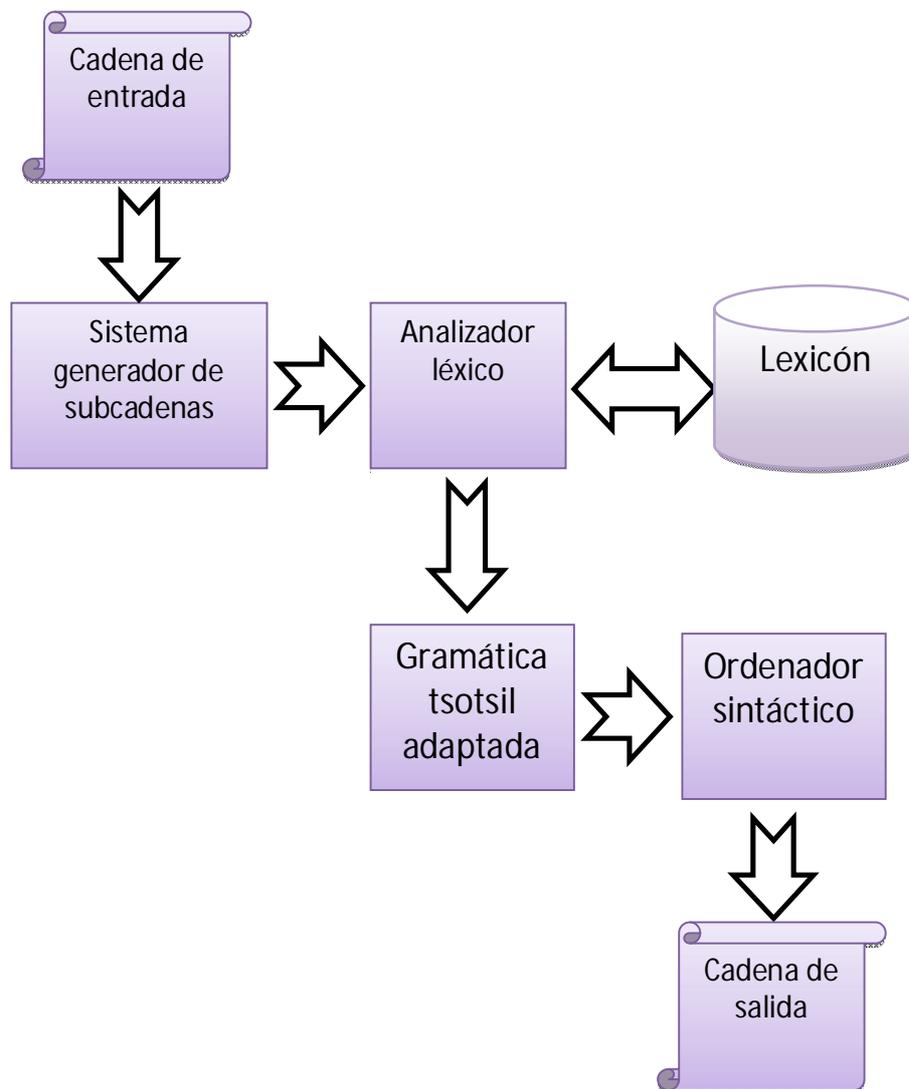


Fig. 1.14 Diagrama de bloques.

### 10.2 Diagrama de secuencia:

Aquí se representa las relaciones entre los objetos y la secuencia de los mensajes del sistema traductor. Podemos observar como se asocian las clases a una operación o a un caso de uso.

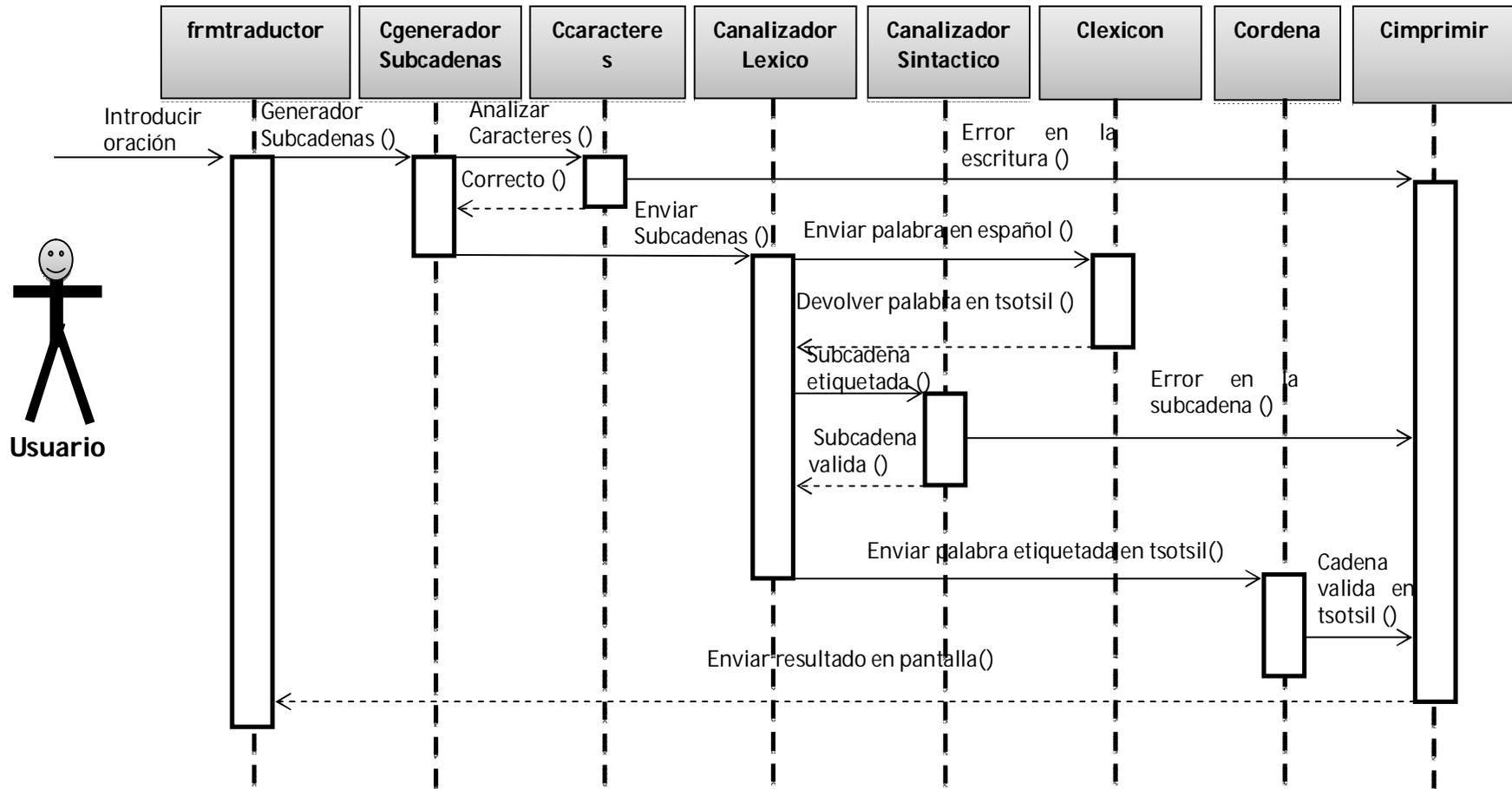


Fig.1.15 Diagrama de secuencia

## 11. CASOS DE PRUEBA

Como se planteó en el objetivo número 1 que Traducirá oraciones simples en tiempo presente del español al tsotsil.

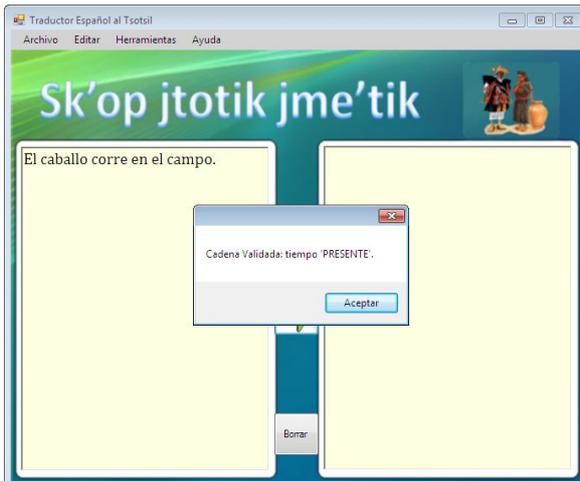


Fig. 1.16 Traducción tiempo presente

El objetivo número 2 es traducir oraciones simples en tiempo pasado:

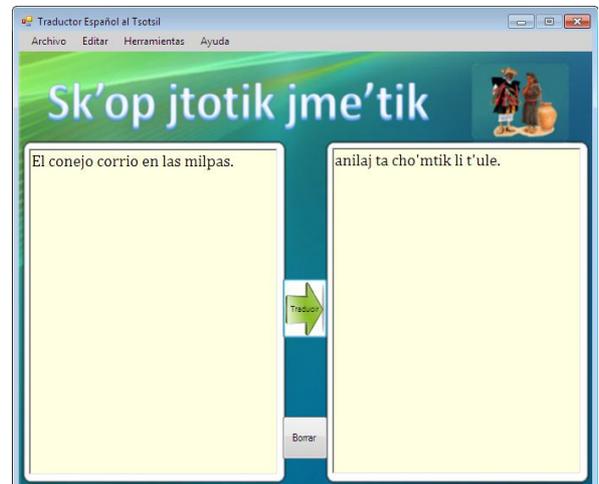
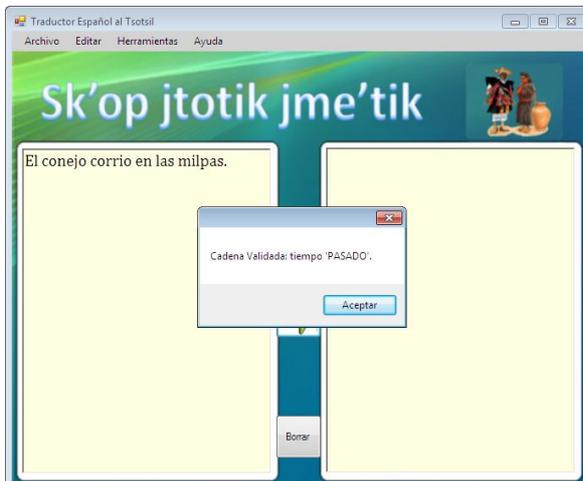


Fig. 1.17 Traducción tiempo pasado.

Otro de los objetivos es que cuente con un diccionario de español-tsotsil.

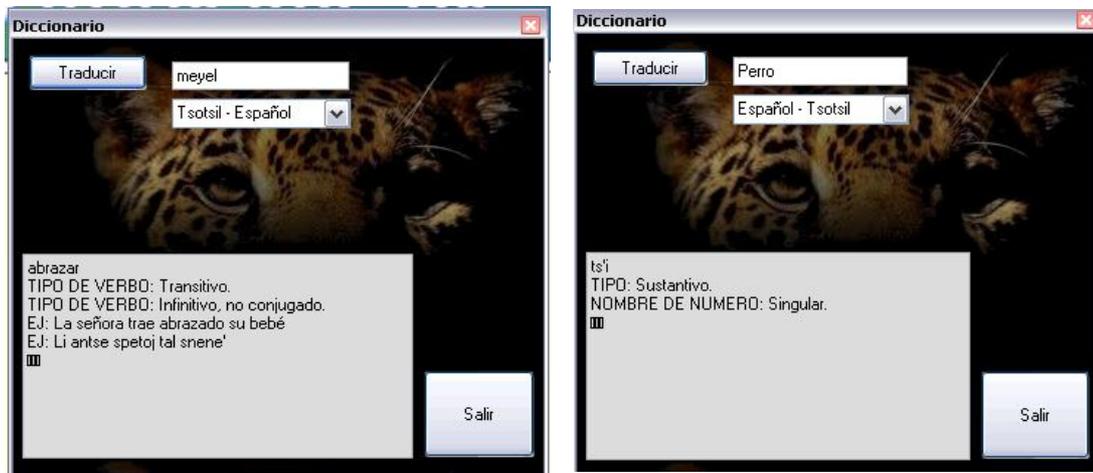


Fig. 1.18 Diccionario español-tsotsil

El objetivo número 4 es imprimir la oración traducida como se muestra en la figura 1.19.

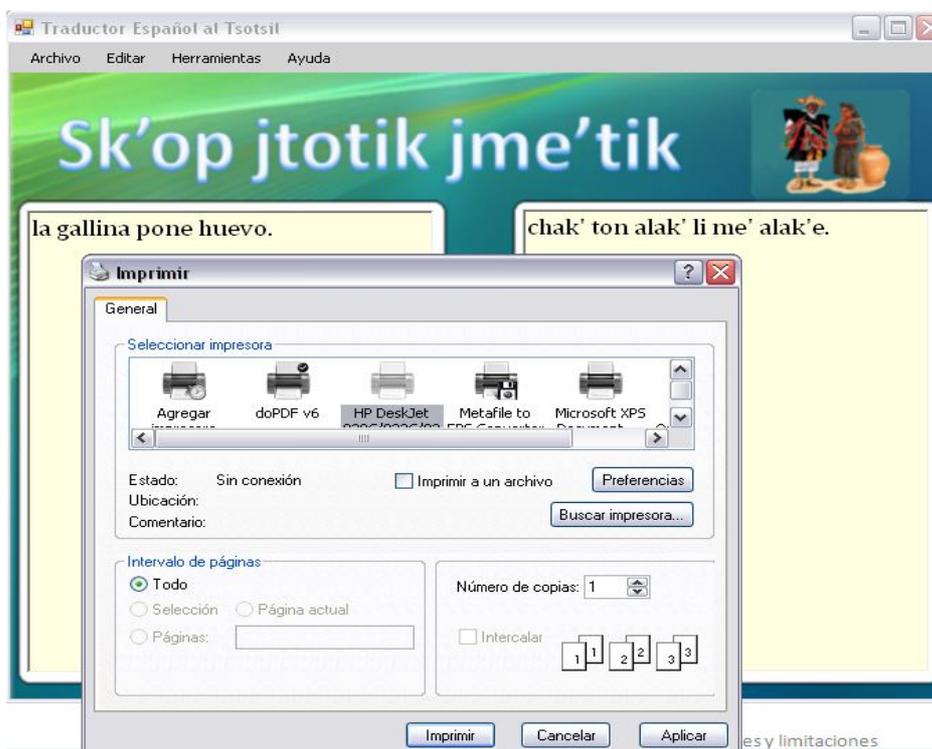


Fig. 1.19 Imprimir oración.

Lo último es guardar la traducción generada en formato \*.txt, \*.pdf o \*.doc. Como se muestra en la figura 1.20.

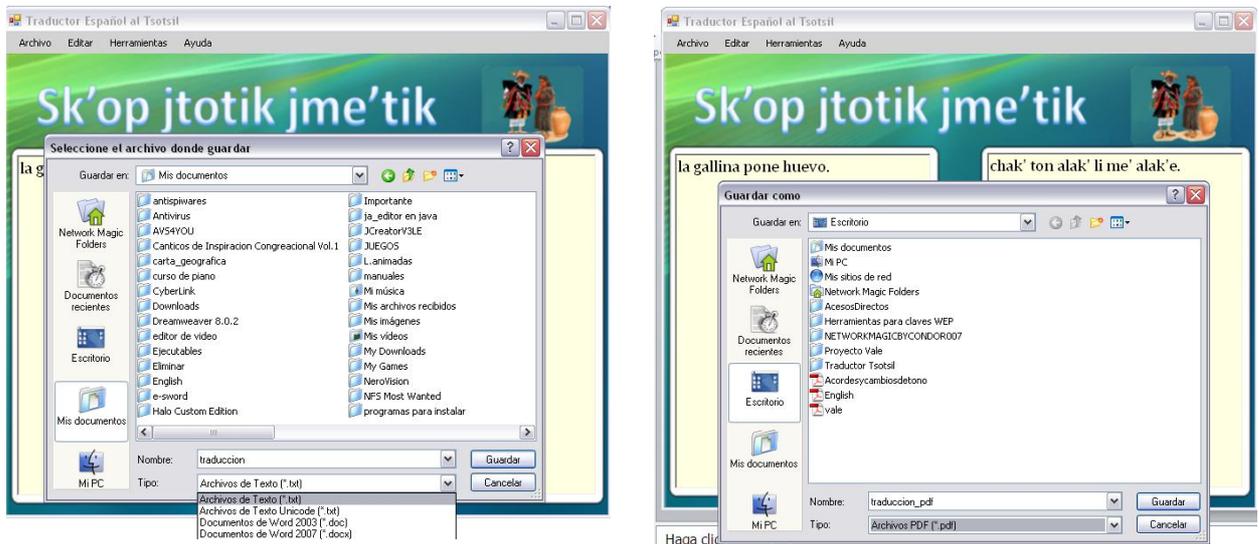


Fig. 1.20 guardar la oración traducida.

También cuenta con una fuente de datos, para cambiar el estilo de letra y el tamaño como se muestra en la siguiente figura 1.21.



Fig. 1.20 cambiar formato fuente y estilo de letra.

## 12. CONCLUSIÓN Y TRABAJOS FUTUROS

Como se observo en los casos de prueba se lograron los objetivos planteados, la construcción de las gramáticas en español y en tsotsil fueron apegadas a las reglas gramaticales ya que el sistema traductor puede traducir oraciones simples en tiempo presente de español a tsotsil y oraciones simples en tiempo pasado de español a tsotsil Por lo tanto observando los resultados de las pruebas realizadas al traductor tsotsil se concluye que los objetivos específicos propuestos para la elaboración de este Proyecto fueron realizados satisfactoriamente.

### Trabajos futuros

Como se pudo observar, el Proyecto del sistema traductor solamente contempla dos tipos de oraciones simples que son en tiempo presente y en tiempo pasado, por lo que se pretende que se realice la traducción de oraciones simples en tiempo futuro y oraciones compuestas en español a tsotsil.

### 13. BIBLIOGRAFÍA

[1] *II Censo de Población y Vivienda 2005. Perfil sociodemográfico de Chiapas*. INEGI: Instituto Nacional de Estadística Geográfica e Informática. México, 2008.

[2] LÓPEZ Gómez, Josías. *Pensamiento y palabra*. Consejo estatal para la cultura y las artes de Chiapas. Tuxtla Gutiérrez, Chiapas, 2003.

[3] ROBLES Vásquez, Héctor V. (coordinador). *Panorama Educativo de México 2007. Indicadores del Sistema Educativo Nacional*. Instituto Nacional para la Evaluación de la Educación. D.F. 2007.

[4] ROBLES Vásquez, Héctor V. (coordinador). *Panorama Educativo de México 2008. Indicadores del Sistema Educativo Nacional*. Instituto Nacional para la Evaluación de la Educación. D.F. 2008.

[5] <http://www.lefalme.selfip.com/spe>

[6] <http://consultacatalogo.educa.madrid.org>

[7] <http://www.encyclomedia.edu.mx>

[8] <http://memnon.ii.uam.es/~educanix/>

[9] <http://www.pipoclub.com/>

[10] [http://dev.laptop.org/~edsiper/byteofpython\\_spanish/ch01s02.html](http://dev.laptop.org/~edsiper/byteofpython_spanish/ch01s02.html)

[11] <http://petra.euitio.uniovi.es/~i1668308/apuntes/lenguaje%20C%2B%2B.pdf>

[12] <http://aulex.org/nah-es/>

[13] <http://portal.perueduca.edu.pe/webtvpublico.htm>

[14] [http://www.worldlingo.com/es/products\\_services/worldlingo\\_translator.html](http://www.worldlingo.com/es/products_services/worldlingo_translator.html)

[15] <http://diccionarios.educacionchiapas.gob.mx/busqueda/buscar2>

[16] *La oración simple y la oración compuesta*. Wikillera to. Disponible en forma digital en: [http://portales.educared.net/wikiEducared/index.php?title=La\\_oraci%C3%B3n\\_simple\\_y\\_la\\_oraci%C3%B3n\\_compuesta&oldid=17681](http://portales.educared.net/wikiEducared/index.php?title=La_oraci%C3%B3n_simple_y_la_oraci%C3%B3n_compuesta&oldid=17681), consultado en septiembre de 2009.

[17] *Ciclo de vida del software. Implementación y debugging.- Métodos del ciclo de vida*.