



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

“DISEÑO Y DESARROLLO DE UNA INFRAESTRUCTURA COMPUTACIONAL BÁSICA PARA EL APRENDIZAJE DEL MAPUZUGUN”

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS, MENCIÓN COMPUTACIÓN

CRISTIAN EDUARDO AHUMADA OLIVA

PROFESOR GUÍA:
CLAUDIO GUTIÉRREZ GALLARDO

MIEMBROS DE LA COMISIÓN:
DORIS SÁEZ HUEICHAPAN
BÁRBARA POBLETE LABRA
ANTONIS ANASTASOPOULUS

SANTIAGO DE CHILE
2022

RESUMEN DE LA TESIS PARA OPTAR AL GRADO
DE: Magíster en Ciencias, Mención Computación
POR: Cristian Eduardo Ahumada Oliva
FECHA: 2022
PROFESOR GUÍA: Claudio Gutierrez Gallardo

DISEÑO Y DESARROLLO DE UNA INFRAESTRUCTURA COMPUTACIONAL BÁSICA PARA EL APRENDIZAJE DEL MAPUZUGUN

El Mapuzugun es la lengua del Pueblo Mapuche y, según la literatura, en el siglo XVI contaba con aproximadamente un millón de hablantes. Luego de esto, ha tenido una fuerte caída en su cantidad de hablantes debido a razones políticas e históricas, donde tanto el Imperio Español como los Estados Chileno y Argentino han contribuido de manera activa en la exclusión de este idioma de los espacios públicos y la educación, quedando relegado principalmente a núcleos familiares y a ceremonias Mapuche en comunidades. Este idioma enfrenta exclusión de medios impresos y, junto a esto, existe una diversidad importante de formas de escribir el Mapuzugun, sin tener una estandarización validada por todos los hablantes. Sumado a que por su tipología, aglutinante y polisintética, no todas las herramientas actuales de idiomas que hoy son dominantes son aplicables directamente a la lengua Mapuche.

Así como su proceso de minorización no fue por causas naturales, hoy el poder revitalizar el Mapuzugun en todos los espacios de la sociedad depende de diversas iniciativas de grupos y organizaciones. Por esto, se enseña y se aprende esta lengua en gran parte del territorio controlado por el Estado de Chile. Dentro de estos espacios e iniciativas también se debe pensar en los espacios relacionados con Internet y la computación, donde aún son pocas las propuestas para aportar en la revitalización del Mapuzugun. Esta tesis se inscribe en ese marco.

La tesis propuso y desarrolló tres módulos para apoyar el aprendizaje del Mapuzugun: Un “detector de grafemarios”, que indique cuál puede ser el sistema de escritura de un texto en Mapuzugun, un “analizador morfológico”, que separa una palabra en su composición interna, y un “traductor informal”, que asigna traducciones al castellano a estas partes o sus combinaciones. Estos módulos fueron integrados en línea y mostrados a los usuarios a través de una interfaz web. Estas herramientas fueron probadas con alrededor de 40 personas relacionadas con el Mapuzugun a través de encuestas y pruebas del software.

Se puede ver que este desarrollo cumple su función de aportar en la aprendizaje del Mapuzugun, ayudando en traducciones de palabras y frases pero, además, se ve que podría ser útil para aprovechar las funciones de este software el tener un nivel a lo menos intermedio en el idioma. Este software, como toda herramienta del lenguaje, se puede (y se seguirá) trabajando. Queda como producto de esta tesis el software actual a disposición pública, síntesis del aprendizaje y enseñanzas que dejó su desarrollo, y propuestas de trabajo futuro relacionados a este software y al Mapuzugun en general.

Dedicatoria

*Zewmafiñ tüfachi küzaw
witxañpüramtuaſiel tati Mapuzugun
welu ka zewmafiñ tukulpayafiel
kom tati pu Mapuzugufe em,
pu Mapuzugufe metu amutunolu,
tati pu Mapuzugufe metu lleqnolu
ka kom tati pu chillkatupelu
ka kimeltupelu tüfachi zugun.
Iñchiñ Mapuzuguliyiñ
taiñ pu pichikeche ka Mapuzugutuaigün.
Femgechi Kom iñchiñ
zoy Mapuzuguyekümelleaiñ.*

Agradecimientos.

Iñche ayüken tañi chillkatual Mapuzugun welu müte ayükelan tañi chillkatual kake zugu, fey mu wünelu mu mañumafñ tati Mapuzugun tañi küllfünelketu mu ka tañi newenelketu mu wiñotual universidad mu chillkatual.

Ka mañumafñ tañi kompañ/afkazi/ayün/che/oye Felisa Manquepi Mariluan, fey mülekelu iñchu mür reküluwal tayu küzaw mu ka tayu chillkatun mu. Ka femgechi mür pichi kellu witxañpüramtukeyu tati Mapuzugun, welu zoy mañumafñ tati rüpu tayu mür amuleken mu.

Goymakelan tañi mañumafiel tañi chaw Ricardo Ahumada Barra, tañi ñuke Carmen Oliva Leiva ka tañi peñi Ricardo Ahumada Oliva. Mülenofule egün, kachegeafun. Metu tañi konnon tüfachi chillkatun mu, mülefuy tañi chuchu (Mami) em welu metu tañi aftunon tüfachi zugu amuturki. Fey azkintunarümnekepelaenu chew püle rume mülele fey mu mañumafñ tañi chuchu em, fey ka kimelenu chumgechi zewmayafiel pagkutxa, ka femgechi nüwyekümei tati zugu. Ka mañumafñ tañi chezky em, tañi kuku em, tañi laku em, tañi ñukentu, ñi epu müna ñuke püle ka tañi pu malle, tañi pu palhu ka tañi pu müna chaw püle. Tañi kompañ püle ka mañumafñ kom kizu tañi pu che, fewla iñche tañi pu che getuygün ka, femgechi amulei mogen ka feyegün ka kimelkenu egün kakewme zugu tachi rüpu mu, zoyelchi tañi llalla ka tañi chezküy tañi küme llowketu amukelu fey mu.

Mañumafñ tañi kimeltuchefe Claudio Gutierrez tañi kelluketu tüfachi zugu mu, ka tañi küllfüneletew ñi entual tüfachi küzaw, zoy mülele femgechi kimeltuchefe kaletuafuy tati academia ka fey zoy kellukeafuy tati pu che. Ka mañumafñ tañi kuifi kimeltuchefe Doris Saez, müte pewelafñ welu fey ka kelluenu tañi amuleal tüfachi Mapuche rüpu mu. Claudio ka Doris manelkenu egu entuafiel küzaw kelluafulu tati pu Mapuche zugu. Re fey egu kellulaenu, ka mañumafñ Antonis Anastasopoulos mülelu Estados Unidos mapu mu, Iñaki Alegría mülelu Euskal Herria mu, Andrés Chandía mülelu Catalan Mapu mu ka tachi mapu mu ka mañumafñ Felipe Bravo ka Francisco Gutierrez tañi kelluetu mu tüfachi küzaw mu kom egün.

Wall mapu mu ka mañumafñ tati pu Mapuzuguletuaiñ che fey egün pütxa küzaw entukigün ka kelluenu egün ka. Ka kiñe mañumafñ kom tati pu Mapuzugufe em, pu Mapuzugufe metu amutunolu, tati pu Mapuzugufe metu lleqnolu ka kom tati pu chillkatupelu ka kimeltupelu tüfachi zugun. Ka mülei tañi mañumafiel kom tati llowzuguetu tañi ramtukan

Mapuzugun chew püle rume allkütugeai ka kom eypiai. Kiñe antü topaañ mogeal tüfachi zugun mu, tati kewün taiñ eluetu tati pu kuifikeche miyawkefulu tüfa püle, tati kewün taiñ eluetu tati gen mapu. Kizu günewkelai che, fey mu ta mülepan fachiantü ka rume mañumkülen tañi eluñmaetu tañi küzawal tüfachi zugu mu.

Agradecemos el apoyo y financiamiento del Programa de Pueblos Indígenas, Dirección de Diversidad y Género, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile.

Tabla de contenido.

Resumen	i
Dedicatoria	ii
Agradecimientos.	iii
1 Introducción.	1
1.1 Objetivo General.	2
1.2 Objetivos Específicos.	2
1.3 Preguntas de investigación.	3
1.4 Hipótesis.	3
1.5 Plan de trabajo.	4
1.5.1 Antecedentes y trabajo relacionado.	4
1.5.2 Obtención de datos.	4
1.5.3 Diseño de las herramientas.	4
1.5.4 Implementación del software.	4
1.5.5 Validación del software.	4
1.5.6 Adaptación y evaluación.	4
1.6 Alcances.	5
2 Antecedentes y trabajo relacionado.	6
2.1 Procesamiento de lenguaje natural.	6
2.2.1 Aplicaciones.	7
2.2 Interacción Humano-Computador.	7
2.3 Mapuzugun.	8
2.3.1. Historia de la lengua.	8
2.3.2. Estado actual.	9
2.3.3. Lingüística del Mapuzugun.	10
2.4.5. Trabajos de lingüística computacional sobre Mapuzugun.	11
2.4 Otras lenguas.	13
3. Metodología.	16
3.1. Obtención de datos.	17
3.2 Traductor de grafemarios.	18
3.3 Analizador morfológico.	20
3.4 Traductor informal.	22
3.5 Integración de las herramientas.	23
3.6 Interfaz de usuario.	24

3.7 Prueba con usuarios.	25
4. Traductor de grafemarios.	26
4.1 Reglas de paso de grafemarios.	26
4.1.1 Diferencias de grafemarios.	26
4.1.2 Orden de conversión.	28
4.2 Detector.	30
4.2.1 Posibles variaciones.	31
5 Analizador morfológico.	33
5.1 Desarrollo.	33
5.2 Algoritmo.	34
5.3 Traductor Informal.	35
6. Implementación del software.	38
6.1 Proceso de una palabra.	38
7. Pruebas del software	43
7.1 Encuesta	45
7.1.1 Primera prueba.	45
7.1.2 Segunda Prueba.	48
7.2 Pruebas del apoyo lingüístico	49
7.2.1 Primera Prueba.	49
7.2.2 Segunda Prueba.	49
7.3 Pruebas de usabilidad.	52
7.3.1 Primera prueba.	52
7.3.2 Segunda prueba.	52
7.4 Los comentarios.	55
7.4.1 Primera prueba.	55
7.4.2 Segunda prueba.	57
7.5 Modificaciones hechas a partir de la retroalimentación.	59
8 Conclusiones y Trabajo Futuro.	63
8.1 Conclusiones.	63
8.2 Trabajo futuro.	64
8.2.1 Posibles trabajos directamente relacionados.	64
8.2.2 Trabajo futuro no directamente relacionado.	65
8.2.2 Trabajo futuro de otras lenguas.	65
9. Bibliografía.	67

Anexos.	73
A. Código reglas de conversión entre grafemarios.	73
A.1 Unificado a Ragileo.	73
A.2 Unificado a Azümcheffe.	73
A.3 Ragileo a Azümcheffe.	73
A.4 Ragileo a Unificado.	74
A.5 Azümcheffe a Unificado.	74
A.6 Azümcheffe a Ragileo.	74
B. Orden de los sufijos.	75
C Comentarios de los usuarios.	76
C.1 Primera Prueba:	76
C.2 Segunda Prueba:	84

1 Introducción.

El Mapuzugun es la lengua hablada por el pueblo mapuche, pueblo indígena originario de las zonas centro y sur de los territorios que hoy controlan los Estados Chileno y Argentino. La lengua está consagrada como derecho por la Declaración de las Naciones Unidas sobre los Derechos de los Pueblos Indígenas. El Mapuzugun es objeto de diversos estudios académicos hoy día. [34,35,38]

Hoy el Mapuzugun está considerado en peligro por la UNESCO, debido a razones históricas [1]. La cantidad de hablantes cada vez es menor [2] y de no hacer algo para revertirlo, la lengua se podría perder [3]. Según Tarcila Rivera Zea [4], “Nuestros idiomas son fundamentales para la afirmación de nuestras identidades, para la continuidad de nuestras culturas y cosmovisiones y para la transmisión intergeneracional de nuestros conocimientos. Son una expresión de libre determinación, diversidad y resistencia”. A partir de aquí, se puede entender la importancia de aportar a este tema de forma transversal en la sociedad y desde todas las áreas del conocimiento.

Para revertir la pérdida de una lengua, se deben generar espacios para ella. Estos espacios van desde lo más cotidiano a esferas más públicas en donde se pueden desarrollar las lenguas. En el caso del Mapuzugun, la falta de acceso a estos espacios no tiene que ver con la naturaleza o la capacidad del Mapuzugun mismo, sino que es consecuencia de la historia del pueblo mapuche [1] y las decisiones políticas tomadas por los Estados chileno y argentino. Si hoy no está en diversos espacios públicos, es porque se le ha relegado, principalmente, a espacios familiares y espirituales dentro de comunidades en sectores rurales en su mayoría, negándole su participación en colegios, universidades, trabajos, instituciones del Estado o vida pública en general.

Las lenguas con pocos hablantes, con poco corpus oral y textual disponibles, etc. se conocen como “lenguas minorizadas” [40,41,52] , entre las que está el Mapuzugun. Hoy día existen diversos esfuerzos para revitalizarlas. En este proceso, ha sido muy relevante el apoyo de herramientas computacionales, desde el proceso de recolección y procesamiento de corpus, de aplicación de técnicas de lenguaje natural, de plataformas para facilitar su uso y aprendizaje por parte de nuevos usuarios, para desarrollar corpus, etc. Hay una rica experiencia internacional con, y material sobre, lenguas minorizadas en lugares como por ejemplo Perú, México, Nueva Zelanda y Euskal Herria 1 . Respecto del Mapuzugun, hoy existen diversas iniciativas para el aprendizaje y la enseñanza con las cuales se busca revitalizar la lengua dentro de distintos espacios, principalmente, educativos.

En el marco de lo expuesto, el objetivo de esta tesis apunta al proceso de revitalización del Mapuzugun.

El objetivo será mostrar que es posible diseñar y desarrollar herramientas computacionales básicas para trabajar en el área del procesamiento del lenguaje natural del Mapuzugun, así como para el aprendizaje de la lengua. Esto permitirá, por un lado, apoyar los esfuerzos que se hacen por revitalizar el Mapuzugun por medio de la enseñanza y la promoción, y por otro, aportar a un campo de investigación en lingüística computacional que conocido como lenguas minorizadas, y en particular, a la lingüística computacional del Mapuzugun.

1.1 Objetivo General.

El objetivo será desarrollar un software simple que permita al estudiante inicial de Mapuzugun consultar palabras y frases, tanto para obtener su significado como para su análisis morfológico basado en herramientas computacionales básicas para el procesamiento lingüístico de esta lengua. Este software estará orientado a estudiantes (particularmente iniciales) de la lengua. Se espera además que, parte de este desarrollo, sirva de base para el desarrollo de herramientas más complejas, tanto para el ámbito de la lingüística computacional del Mapuzugun, como el de la revitalización lingüística referido al aprendizaje y enseñanza de esta lengua.

1.2 Objetivos Específicos.

1. Diseñar e implementar un traductor inter-grafemarios del Mapuzugun, para los tres grafemarios más reconocidos, tomando en cuenta la inconsistencia de los textos.
2. Adaptar el analizador morfológico para el Mapuzugun, existente en [8], que pueda segmentar las palabras, de forma lingüística, para su uso en espacios de enseñanzas, filtrando aspectos para el entendimiento de personas no expertas en lingüística.
3. Desarrollar e implementar a partir de analizadores morfológicos existentes, un traductor informal, esto es, que dé al estudiante nociones e intuiciones de cómo se podría traducir formalmente una cierta palabra y/ o frase del Mapuzugun.
4. Validar las herramientas básicas (1)-(3) desarrolladas, con expertos en la lengua y con estudiantes de Mapuzugun.
5. Integrar el traductor inter-grafemarios, el analizador morfológico modificado y el traductor informal en un proceso que muestre el análisis de palabras y frases
6. Generar, a partir de las herramientas validadas, un software simple que pueda ser usado por quienes enseñan y/o estudian el Mapuzugun, junto con una prueba de usabilidad y un estudio de su impacto en la enseñanza y aprendizaje.

1.3 Preguntas de investigación.

1. ¿Es posible desarrollar herramientas lingüísticas computacionales más formales para una lengua que hoy tiene diversidad de hablantes, diversidad de grafemarios y escasos corpus formales?
2. ¿Qué herramientas básicas de Procesamiento de Lenguaje Natural, dado el contexto actual del Mapuzugun, se requerirían para generar aplicaciones más complejas de PLN como Traducciones automáticas, NER, etc?
3. ¿Qué herramientas básicas del área de procesamiento de lenguaje natural se necesitan, dado el contexto actual del Mapuzugun, para el apoyo a su aprendizaje y enseñanza?
4. ¿Cuáles de estas herramientas se pueden desarrollar dado el contexto actual del Mapuzugun y sus recursos disponibles?
5. ¿Qué recursos se necesitan para poder desarrollar las herramientas que todavía no son alcanzables en un corto plazo y que contribuyan al desarrollo de la lingüística computacional del Mapuzugun?
6. ¿Es posible desarrollar un software que impacte positivamente en el aprendizaje del Mapuzugun?

1.4 Hipótesis.

La siguiente es la hipótesis de la tesis:

Es posible, dado el estado del Mapuzugun, desarrollar y disponibilizar un software simple para el Mapuzugun que:

- (a) aporte a la revitalización del Mapuzugun impactando en el ámbito de la enseñanza de esta como segunda lengua, mejorando el aprendizaje de los estudiantes de esta lengua;
- (b) contribuya a futuros investigadores y desarrolladores de aplicaciones de PLN de Mapuzugun.

La tesis pretende probar esta hipótesis por medio del diseño, desarrollo e implementación del software. Se plantea probar (a) con distintos estudiantes y profesores de Mapuzugun de Santiago y Temuco, principalmente en sus aspectos de usabilidad del software y ayuda al proceso de aprendizaje de la lengua. Para (b) se pretende mostrar que con esas herramientas se pueden generar y mejorar corpus existentes.

1.5 Plan de trabajo.

A continuación se presentan los pasos para poder cumplir los objetivos planteados:

1.5.1 Antecedentes y trabajo relacionado.

Recopilación y análisis de trabajos relacionados al procesamiento de lenguaje natural, la interacción humano-computador, el Mapuzugun y literatura relacionada a otras lenguas minorizadas en el contexto computacional.

1.5.2 Obtención de datos.

Hecha la recopilación de antecedentes, se procede a obtener datos que serán relevantes para la validación de las herramientas, como diccionarios y extracción de palabras en mapuzugun de distintas fuentes para probar las segmentaciones y detección de grafemario.

1.5.3 Diseño de las herramientas.

Se generan las ideas de los algoritmos para las herramientas, y su forma de integración para lograr el software que se plantea. Además se plantean las decisiones para la implementación con usuarios.

1.5.4 Implementación del software.

Se construyen las herramientas por separado y se validan para, luego de terminadas, hacer una integración de estas como conjunto detección-segmentación. Finalmente, se adapta el conjunto para ser utilizado por usuarios usuario tomando las herramientas y poniéndolas bajo una interfaz.

1.5.5 Validación del software.

Se prueba el software con distintas personas que estudian y enseñan Mapuzugun con el fin de validar tanto la parte lingüística (segmentaciones y detecciones de grafemario) como la parte de usabilidad. Se analizarán estos resultados para ver qué cambios se necesitan hacer al software para que sea más usable, no tenga errores lingüísticos y sirva como apoyo a la enseñanza y aprendizaje del Mapuzugun.

1.5.6 Adaptación y evaluación.

A partir de la evaluación, se modifica el software para que se acerque más a lo que esperan los usuarios, de ser necesario podría haber más de un paso de validación-adaptación, dependiendo de los comentarios y la cantidad de cambios a hacer.

1.6 Alcances.

El software a construir tendrá la capacidad de funcionar segmentando palabras de Mapuzugun, una vez detectado el grafemario en que están escritas las palabras que se introducen. Esta dará todas las combinaciones posibles bajo las reglas gramaticales del Mapuzugun. Este software quedará disponible de forma online para que sea ocupado por usuarios de internet interesados en el Mapuzugun. Además, este software quedará liberado para que las personas que enseñan y aprenden lo puedan usar sin problemas. Finalmente se entrega la experiencia del desarrollo de este software y los futuros desarrollos que podrán enriquecerlo y mejorarlo.

1.7 Organización de la Tesis.

El contenido de este trabajo está organizado de la siguiente forma: Este primer capítulo introductorio, luego le sigue un capítulo de “Antecedentes y trabajo relacionado”, en donde se presentan artículos y experiencias de softwares relevantes o relacionados con esta tesis. En el capítulo tercero, se aborda la metodología y el cómo se abordarán las unidades del software y su integración. Luego, el capítulo cuarto explica la forma de implementación del traductor de grafemarios y su uso como un detector de grafemarios para palabras en Mapuzugun. De forma similar al cuarto, el capítulo quinto explica la implementación del Analizador morfológico, el cual separa palabras en mapuzugun. El capítulo sexto contiene la forma de implementación del software y, además, se habla de la integración de las unidades de las herramientas de los capítulos cuarto y quinto, y la construcción de la interfaz para que puedan ser usada por usuarios. El capítulo séptimo corresponde a los resultados de las pruebas hechas con usuarios. Y, finalmente, el octavo habla de las conclusiones obtenidas de esta tesis y de cómo seguir en el futuro.

2 Antecedentes y trabajo relacionado.

Revisaremos cuatro aspectos que tienen relevancia con el tema. Primero, una visión general de PLN y la Interacción Humano-Computador. Luego el estado de los trabajos de lingüística computacional de Mapuzugun. Finalmente, trabajos sobre otras lenguas minorizadas.

2.1 Procesamiento de lenguaje natural.

El Procesamiento de Lenguaje Natural (PLN) es un área de la computación que se encarga de darle capacidad a los computadores de hacer tareas útiles con los lenguajes humanos. Es un área desafiante debido a que estos lenguajes son complejos y llenos de ambigüedades para las máquinas. Estas tareas se refieren principalmente a permitir una buena comunicación humano-computador o mejoras en la comunicación humano-humano [5].

A partir de estas premisas y las necesidades de los distintos usuarios y sus idiomas, surgen una serie de herramientas y aplicaciones para resolver los problemas que se plantean. Es posible distinguir distintos niveles del lenguaje, que podríamos sintetizar en los siguientes [6]:

1. *Fonético*: En el nivel fonético, se estudian los sonidos del lenguaje humano.
2. *Fonológico*: Este nivel estudia los distintos sonidos agrupados en sistemas que podrían corresponder a ciertos idiomas.
3. *Morfológico*: Análisis de las unidades mínimas del lenguaje, como podrían ser palabras, sufijos, prefijos, entre otros, que componen las palabras o frases del lenguaje.
4. *Sintáctico*: Se refiere al nivel en el cual se analiza el ordenamiento de las unidades básicas del lenguaje, entrando en las estructuras de ciertas palabras o frases.
5. *Semántico*: En la semántica se analizan los significados de las expresiones lingüísticas dentro del lenguaje.
6. *Pragmático*: Está relacionado con el uso de las expresiones, que comprende ya no solo la palabra, si no que, además su contexto en la situación particular a estudiar.

A partir de estos niveles se pueden agrupar ciertas aplicaciones del PLN, que podrían contener elementos de más de una de estas categorías.

2.2.1 Aplicaciones.

Dentro del PLN existen ciertas tareas generales que se pueden desarrollar para los distintos lenguajes a analizar. Algunas de ellas son:

1. Clasificación de tópicos.
2. Análisis de sentimiento.
3. Traducción automática.
4. Extracción de información.
5. Sistemas de interacción humano-computador.
6. Generación de texto.
7. Reconocimiento de voz.
8. Generación de voz a partir de texto.
9. Responder preguntas.

Muchas de estas tareas tienen un avance significativo pero, generalmente, sobre idiomas muy bien establecidos (en términos de su formalización, gramática, léxico, etc.) y hablados por grupos sociales dominantes y, principalmente, en inglés, debido a que es uno de los idiomas más utilizados del mundo y de los que tienen la mayor cantidad de textos o un corpus más grande. En contraste con esto, hay idiomas que, aunque ampliamente hablados y establecidos, no tienen la magnitud de datos ni esfuerzo invertido del inglés, como por ejemplo, el idioma turco, que es un idioma oficial de un Estado importante. Y luego están las lenguas como el Mapuzugun, las lenguas originarias de Latinoamérica y muchas otras lenguas, que han sido minorizadas en las distintas regiones del mundo, negándoles o quitándoles su estatus al ignorarlos en leyes y programas oficiales. [40,41]

2.2 Interacción Humano-Computador.

La interacción humano-computador es el área que se encarga del estudio de la manera en que los usuarios interactúan con los computadores. Es importante en este trabajo, debido a que es un software y necesita tener una usabilidad e interfaces pertinentes. El caso más común puede ser el cómo se percibe la interfaz de un software, en nuestro caso para el aprendizaje del mapuzugun. En este dominio se pueden medir criterios importantes para que el software, en términos sencillos, cumpla su objetivo, que es apoyar el aprendizaje del idioma. Estos criterios según Andreas Holzinger [55] son la capacidad de aprendizaje, eficiencia memorabilidad, tasa de error baja y satisfacción. A partir de estos criterios, se pueden separar distintos métodos entre los que encontramos los que consideran usuarios y los que no lo hacen, llamados “Métodos de inspección” y “Métodos de pruebas de usabilidad”, respectivamente. Además existen otros métodos en la literatura que se refieren a otros aspectos, más generales y que agregan criterios que solo lo técnico [56].

2.3 Mapuzugun.

En esta sección, para entender el contexto y la necesidad de este trabajo, se explicará cómo el Mapuzugun pasó de ser una lengua de un millón de hablantes en el siglo XVI a convertirse, según la Unesco, en una lengua en peligro en estos días. Se detalla el estado actual de la lengua y como se define de forma lingüística, además de los trabajos computacionales existentes sobre el Mapuzugun.

Para el caso de la denominación del Mapuzugun, existe una extensa discusión en la academia sobre las diferencias entre los términos idioma, lengua y dialecto, algunas definiciones de estos términos se pueden consultar en [53][54]. A partir de esto, en este documento elegimos “lengua” por ser el término más general, aunque para ciertos casos pueden ser intercambiable con “idioma”, y aun en otros ámbitos se prefiere el término idioma para posicionar el Mapuzugun y darle un estatus, en comparación con los idiomas o lenguas dominantes.

2.3.1. Historia de la lengua.

El pueblo mapuche tiene su origen en el territorio conocido como Wallmapu (lo que se puede considerar como el País Mapuche [1]). Este territorio abarca desde Coquimbo hasta Chiloé, incluyendo además zonas del "otro lado" de la cordillera, como Neuquén, en una vasta zona demarcada por el Río Negro. En todo este territorio se pueden encontrar distintas denominaciones para este pueblo que comparte muchos aspectos culturales [1]. Este es el ámbito en el cual se puede enmarcar --dentro de lo que los españoles definían a finales del siglo XVI-- los alcances de la lengua conocida como Mapuzugun.

El Mapuzugun, en toda su extensión, a la llegada de los españoles, era hablado por alrededor de un millón de personas [1]. De aquí su importancia y su contraste con la realidad actual. No por nada uno de los primeros libros que se tiene sobre esta lengua se titula *"Arte y gramatica de la lengva general que corre en todo el Reyno de Chile, con vn Vocabulario y Confessionario"* publicado en 1606 por Luis de Valdivia [28]. Además se puede ver en la toponimia del país lugares con un claro origen del Mapuzugun, como Huentelauquén en el norte, Puchuncaví, Curacaví, Pudahuel, Vitacura, con ejemplos incluso en Puel Mapu (o lo que hoy conocemos como Argentina), hasta, Chiloé en el sur.

Durante la interacción de mapuche con españoles durante la Colonia, se puede apreciar el lugar del Mapuzugun en todas las esferas de la sociedad, desde la familia, hasta relaciones políticas internacionales, como lo fueron los Koyagtun (o Parlamentos) principalmente con el Imperio Español, los Estados Chileno y Argentino, pero también con Franceses, Holandeses, Ingleses. En todos estos destacaba la figura del *linguaraz*, que actuaba como traductor para tratar de reproducir fielmente las ideas que se tenían en Mapuzugun a los representantes extranjeros.

Fue durante la construcción de los Estados nacionales chileno y argentino en el siglo XIX --que en principio no incluían territorio mapuche-- con las llamadas "Campaña del Desierto" y "Pacificación de la Araucanía", cuando estos Estados sometieron políticamente a este pueblo. Junto con ello, como parte de la construcción de las identidades de Chile y Argentina, se fue quitando espacio al Mapuzugun y a la cultura mapuche por medio de las escuelas y la Iglesia, las cuales, a través de la evangelización y los castigos, fueron negando la identidad indígena junto con su lengua.

Luego a principios del siglo XX, después de ese despojo territorial, hubo una fuerte migración mapuche hasta las ciudades más importantes de Chile, en busca de mejores condiciones de vida. Esto se tradujo en pérdidas culturales, muchas veces por racismo y discriminación. Sin embargo, se desarrollaban algunos esfuerzos por parte de los propios mapuches para mantener la cultura y la lengua, como muestran las publicaciones como las de Pascual Coña [29] y Manuel Mankilef [30,31], que estaban escritas tanto en Mapuzugun como en Castellano.

A finales del siglo XX, debido a la migración comentada, ya había una gran cantidad de personas mapuche en las grandes ciudades. Sin embargo, estas personas estaban desconectados de su cultura debido al contexto. Durante la dictadura y desde los años 90 's el pueblo Mapuche comenzó a tener un mayor posicionamiento político. Con ello se fue recuperando la lengua mapuche de la mano de una recuperación de la identidad en diversos ámbitos, además de mantenerse territorios en los que el Mapuzugun se habla como primera lengua. Hoy en día, según el censo 2017, la mayoría de la población Mapuche estaría en Santiago, pero la mayoría no habla ni tampoco entiende su lengua.

Hoy en día, existen diversas organizaciones que hacen cursos o herramientas que aportan a la revitalización del Mapuzugun. Estas instancias tienen un hito en una marcha que se organiza durante febrero, en el marco de la conmemoración del "Día Internacional de la Lengua Materna" [32], teniendo, como movimiento, demandas importantes como la oficialización del Mapuzugun [33].

2.3.2. Estado actual.

Diversas fuentes, estiman entre 100.000 y 300.000 [2,9] la cantidad de hablantes de Mapuzugun. Ellos constituyen cerca de un ~5 o ~5% de la población mapuche (1.745.147), que conforman el 9,9% de la población total de Chile (17.574.003).

Según la Unesco, una lengua está en peligro cuando se deja de utilizar, cuando se usa en menos ámbitos y cuando se deja de transmitir. A partir de esto se plantea que "cerca del 90 % de todas las lenguas podrían ser sustituidas por las dominantes de aquí a finales del Siglo XXI". Todo esto,

sumado a una documentación insuficiente, genera que haya lenguas extintas o en vías de extinción, que son irrecuperables [6].

Existen seis grados para definir el estado de peligro de estas lenguas. Dentro de esta clasificación se encuentra el Mapuzugun, referido como Huilliche y, tanto en Chile como en Argentina, como Mapuche, como se puede ver en el Atlas de la Unesco.[3]. Están en grado 1 (“En situación crítica”, Huilliche), 2 (“Seriamente en peligro o amenazada”, Mapuche, Argentina) y 3 (“Claramente en peligro o amenazada”, Mapuche, Chile).

Pero no solo a través de la Unesco, se han desarrollado investigaciones sobre el estado del Mapuzugun. También existen diversos estudios desde el área de la sociolingüística para entender ciertos procesos actuales de la lengua y su incorporación en las políticas públicas [33,41,42,44,45,46].

2.3.3. Lingüística del Mapuzugun.

Lingüísticamente se define al Mapuzugun como una lengua aglutinante y polisintética, lo que significa que tiene sus expresiones tienen una raíz principal a la que se le agregan sufijos definidos y distinguibles, para formar frases. Por ejemplo la palabra Kim mapuzuguyekümelleaiñ, que se explica en la Tabla 2.1. Ejemplos como Inglés, Chino o Castellano, no están en esta categoría, y por lo tanto las técnicas de procesamiento que sirven en esas lenguas difieren de las técnicas que se podrían usar para el Mapuzugun.

Palabra	Kim mapuzuguyekümelleaiñ
Separación	Kim mapu-zugu-yekü-me-lle-a-iñ
Traducción	Nosotros, sí, iremos aprendiendo la lengua mapuche

Tabla 2.1. Explicación de la aglutunación de la palabra Kim Mapuzuguyekümelleaiñ

Antes de la colonización, el Mapuzugun se describe como una lengua puramente oral. Hoy día, hasta hace poco tiempo no se enseñaba ni usaba formalmente por instituciones públicas y de educación en Chile. Esto ha significado que no cuente con una estandarización de su escritura ni ortografía. Hoy día existen distintas formas de escribirlo y además, variaciones ortográficas territoriales, debido a que en distintas regiones hay diferencias fonéticas para ciertos sonidos y eso se traduce, en general, en escrituras distintas. Hoy existen tres grafemarios principales para escribir el Mapuzugun: Ragileo, creado por Anselmo Ragüileo en 1985; Unificado, creado por María Catrileo en 1989; y Azümchefe, creado por Necul Painemal para la CONADI (Corporación Nacional de Desarrollo Indígena) en 2008.

Entre estos grafemarios se pueden notar ciertas diferencias visibles en la Tabla 2.2. En el caso del Ragileo, este grafemario usa solo un grafema por fonema, y en ciertas ocasiones, los sonidos asociados a estos grafemas no se corresponden con los del castellano, por lo que presenta un poco más de dificultad para ser aprendido que los otros. El Unificado tiene una escritura más parecida al castellano. Aunque, si bien, la mayoría los grafemas son los mismos, hay fonemas que pueden considerarse parecidos, pero que no son iguales entre castellano y Mapuzugun. Finalmente, el Azümchefe es una suerte de punto intermedio, pero también presenta dificultades y diferencias entre grafemas y fonemas del castellano. Es el usado por instituciones públicas como la CONADI.

Esta falta de estandarización del Mapuzugun trae complicaciones a personas que están estudiando la lengua y que solo dominan un grafemario. Esto también incide en el procesamiento del Mapuzugun, pues se tendrían inconsistencias al tomar datos desde fuentes distintas o incluso desde la misma fuente, sobre todo en temas como traducción automática o análisis semántico, donde la misma palabra podría tener variadas formas y afectar el aprendizaje de algún modelo. Probablemente esto incide en la actual falta de herramientas básicas en esta lengua.

En esta dirección, actualmente existen diversos trabajos relacionados puramente a la lingüística del Mapuzugun: descripciones [34,35, 39], pero también artículos académicos específicos de aspectos técnicos de la lengua [38,46,47,48,49,50,51] y diccionarios [42].

2.4.5. Trabajos de lingüística computacional sobre Mapuzugun.

Hoy existen diversas iniciativas de lingüística computacional sobre Mapuzugun. Existe un normalizador ortográfico y un analizador morfológico [8], pero todavía cuenta con algunos errores derivados de que aplica directamente una serie de reglas sin analizar la entrada que recibe. Estos son aspectos que se podrían mejorar. Otro aspecto que se podría mejorar es que, actualmente, no existe la posibilidad de elegir el grafemario de salida, restringiendo a solo una forma de escritura. Esto es inconveniente hoy que aún no existe un acuerdo en la estandarización de la escritura. Esta implementación fue hecha a partir de un conjunto de reglas a través de expresiones regulares, con un transductor de estados finitos, que han sido liberadas en la página web del autor.¹ Este autor trabaja, además, en un prototipo de analizador morfológico y en corrector ortográfico, basados en las herramientas de estados finitos de Xerox. También hay interfaces de explotación de corpus anotadas con estas mismas herramientas, creadas en un máster interuniversitario en Barcelona, coordinado por la Universidad de Barcelona, y se está apuntando a un traductor automático de Mapuzugun. Por encontrarse en proceso de trabajo doctoral, los resultados de estas herramientas aún no han sido publicadas, pero se pueden revisar en su propuesta de tesis [10].

¹ www.chandia.net

Unificado	Ragileo	Azümchefe
A a	A a	A a
Ch ch	C c	Ch ch
D d	Z z	Z z
E e	E e	E e
F f	F f	F f
G g	Q q	Q q
I i	I i	I i
K k	K k	K k
L l	L l	L l
LI ll	J j	LI ll
M m	M m	M m
N n	N n	N n
Ñ ñ	Ñ ñ	Ñ ñ
N n	H h	Nh nh
Ng ng	G g	G g
O o	O o	O o
P p	P p	P p
R r	R r	R r
S s	S s	S s * Sh sh
T t	T t	T t
Tr tr	X x	Tx tx
T t	T t	T' t'
U u	U u	U u
Ü ü	V v	Ü ü
W w	W w	W w
Y y	Y y	Y y

Tabla 2.2. Comparación de grafemarios del Mapuzugun a utilizar en este trabajo (Elaboración propia.)

Por otro lado, existió un proyecto llamado AVENUE, en el cual colaboró la Universidad de la Frontera, el Programa de Educación Intercultural Bilingüe y el Instituto de Tecnologías del Lenguaje de la Universidad Carnegie Mellon (CMU). Ese proyecto tenía como propósitos generar traductores sencillos y de bajo costo, además de ayudar a la preservación de las lenguas indoamericanas. Este proyecto tuvo como resultado cuatro productos:

1. En primer lugar un grafemario para efectos del procesamiento y el desarrollo informático del Proyecto.
2. Un corpus de 170 horas que ha sido transcrito, pero que no se ha revisado completamente.

3. Un prototipo de traducción que consta de un traductor basado en ejemplos, entrenado. Además en paralelo se trabajó uno basado en reglas de transferencia (ambos con el castellano como par). Este prototipo además cuenta con un analizador morfológico. Luego del proyecto Avenue se trabajó también en CMU en el mejoramiento automático de las traducciones.

4. Un corrector ortográfico del que se dice podría contener un estimado de 6.000.000 de palabras, para OpenOffice. Y que consiste en dos diccionarios, uno de raíces y otro de sufijos, que dentro de MySpell de OpenOffice, corrigen un texto de la forma típica a la que está acostumbrado el usuario. Este sigue teniendo la limitación del grafemario, además de no existir la seguridad que al escribir una palabra en otro grafemario éste la convierta al usado por el sistema [11].

En el ámbito educativo, existe un software para aprender Mapuzugun llamado MAPU de un trabajo en Pontificia Universidad Católica de Valparaíso que además incluye reconocimiento de voz para el control de la aplicación, que funciona correctamente, pero no es robusto a la pronunciación [12]. En este trabajo, también hace referencia a otro prototipo de traducción voz texto de Mapuzugun a español, basado en grabaciones y a un chatbot del proyecto Pandora.

Además, el CEDETI de la Pontificia Universidad Católica, que se dedica a trabajar tecnologías para la integración, tiene un software de aprendizaje de la lengua llamado Mapudungun mew [13].

2.4 Otras lenguas.

En esta sección se presenta una serie de documentos, que además de que su contenido se alinea a los objetivos de este trabajo, podrían definir una hoja de ruta en términos de ser un aporte a la revitalización lingüística, junto con la generación de herramientas básicas de procesamiento para lenguajes en peligro, principalmente unidas al Mapuzugun por su contexto pero, además, en algunos casos por su tipología.

En [14] se hizo una revisión de los desafíos para las lenguas indígenas en América en cuanto a tecnologías para el lenguaje y NLP, que además es una revisión de las experiencias que se han tenido para distintas lenguas de todo el continente. Respecto del Mapuzugun se habla del intento del proyecto AVENUE por generar un traductor de Mapuzugun-castellano del mismo y de los datos que se tienen disponibles. Además aborda lenguas como Quechua, Nahuatl, Wixarika, Shipibo Konibo, Guaraní, entre otras. Los desafíos tienen que ver principalmente con los corpus insuficientes o no muy trabajados, los traductores y analizadores morfológicos. Además se nombran experiencias en las distintas tareas comunes para NLP.

Desde la P. Universidad Católica del Perú se generó una hoja de ruta sobre la “Lingüística computacional para la revitalización y el poliglotismo” [4]. Allí se definen objetivos y

metodologías de cómo abordar este tema. Además, este trabajo clasifica las tecnologías y herramientas básicas junto con su rol en tareas más complejas, definiendo incluso plazos y costos que se debe invertir en la generación de las herramientas. Este trabajo, y todos los relacionados con Quechua, toman una mayor relevancia pensando en que esta lengua tiene una tipología similar.

En [15] se puede ver el proceso más completo de lo que sería el resultado del proyecto AVENUE, que tuvo como producto un traductor Quechua - Español. Esto no pudo ser completado para el caso del Mapuzugun, pero se tiene la metodología con la que se podría continuar. Esto según las propias palabras de los autores, sería un primer paso para la resolución del problema.

Se puede también ver el uso de clasificadores bayesianos y K vecinos más cercanos (k-nearest neighbours, KNN) para la desambiguación en traducción del Quechua [16].

También en el Quechua, el mejoramiento de reconocimiento de morfemas a partir de su comparación con el finlandés, debido a que se tienen estructuras similares, sobre todo en la parte de aglutinación [17].

La cercanía en tipología, también sucede con otras lenguas que están en Perú y el resto del continente, como el Mexicanero, el Nahuatl, Wixarika y Yorem Nokki [18]. O el Mohawk y Plains Cree [19, 63], de más al norte.

En la Universidad de Limerick se desarrolló una tesis sobre un analizador morfológico para el caso del Mohawk. Esto se hace a través de estados finitos y su entrenamiento a partir de los datos de la lengua [20].

Existe también en [61] herramientas de análisis morfológico para Arapaho. En [62] distintas herramientas para el Inuktitut y en [64] para el idioma Seneca.

En [21] se presenta un identificador de lenguas peruanas de bajos recursos. Esto se hace a partir de la construcción de un modelo de espacio vectorial para los lenguajes, a partir de bigramas y trigramas, y una matriz a partir de “frecuencia de término – frecuencia inversa de documento” o (TF-IDF, por sus siglas en inglés). Se clasifica por oraciones y se logran un rendimiento sobre el 96 % en clasificación con *support vector machine*. Esto si bien son buenos resultados, no existe una forma de saber si es que es exactamente el grafemario usado o si es que es solo el más cercano.

En [22] Alva y Oncevay proponen un corrector basado en silabificación y caracteres para un lenguaje peruano aglutinante. Esto se hace con grafos de sílabas y caracteres a partir de modelos extraídos de corpus. Este método propone tres correcciones más cercanas para una palabra mal escrita con métricas de distancias por caracteres, también guardando las correcciones previas. Este método es completo y toma en cuenta las sílabas y caracteres, lo que sería importante en el caso de los grafemarios, que tienen diferencias sutiles, como si fueran errores ortográficos.

Aunque el error puede ser mejorable (76 %), podría ser una solución para el normalizador, si es que se amplía a multi idiomas (o en este caso grafemarios).

Saliendo de América, también encontramos desambiguación morfológica estadística para el Turco [23]. También existen desarrollos de aprendizaje levemente supervisado de morfología para lenguajes de bajos recursos aplicados a isiZulu y Turco, que consta de un algoritmo de extracción de raíces y un sistema de segmentación no supervisado para convertir secuencias de prefijos y sufijos, sufijos y prefijos individuales [24]. En otro trabajo, más general, se puede ver segmentación de idiomas aglutinantes, tomando en cuenta los largos de morfemas, lo cual afecta positivamente en los resultados [25].

UralicNLP [60] desarrolló paquetes de python con herramientas de análisis morfológico para lenguas Urálicas y, en general, sus publicaciones tratan los idiomas como más allá de objetos de estudio para la academia.

A partir de estos antecedentes técnicos (NLP y IHC), del estado de peligro en el que está el Mapuzugun, su historia y las aplicaciones construidas para su análisis o aprendizaje, además de los trabajos que se hacen sobre otros que también han sido minorizados se puede tener un contexto sobre cómo aportar a la enseñanza del Mapuzugun a partir de la computación, considerando experiencias tanto a nivel local como internacional, para desarrollar los pasos de la solución que se planteará.

3. Metodología.

En esta sección se plantea cómo se aborda cada punto del plan de trabajo. Como primer elemento, se tiene una vista general de los pasos a seguir para las herramientas, su integración y el desarrollo de la interfaz de usuario, con sus pruebas correspondientes:

1. Reunir los datos de los diferentes corpus y pre-procesarlos, definiendo para cada tarea un formato de ser necesario, según sus particularidades.
2. Implementación y validación del Traductor de grafemarios:
 - a. Desarrollo, implementación y validación del
 - b. Implementación de traductor básico de grafemarios conocidos a partir de expresiones regulares.
 - c. Validación del traductor básico de grafemarios con textos de prueba, en los grafemarios conocidos.
 - d. Implementación de mejora para textos escritos en grafemarios no conocidos o variantes.
 - e. Validación del traductor básico de grafemarios con textos de prueba escritos en grafemarios no conocidos.
2. Implementación y validación del Analizador morfológico.
 - a. Implementación del analizador morfológico, a partir de gramáticas del Mapuzugun.
 - b. Validación del método comparándolo con datos anotados y hablantes
3. Desarrollo, implementación y validación del traductor informal
 - a. Asignar valores entendibles a las unidades morfológicas del analizador morfológico a partir de un diccionario.
 - b. Validar con textos paralelos y hablantes.
4. Integración de las herramientas para el análisis de palabra y de frase, con el fin de construir el software.
 - a. Integrar el proceso para palabra:
 - i. Identificar grafemario
 - ii. Normalizar grafemario
 - iii. Segmentar palabra
 - iv. Mostrar segmentación
 - b. Integrar proceso de frase:
 - i. Identificar grafemario
 - ii. Normalizar grafemario
 - iii. Segmentar palabra(s)
 - iv. Mostrar segmentación

- v. Traducir cada unidad (morfema o palabra indivisible)
 - vi. Dar traducción informal (no correcta gramaticalmente, necesariamente)
5. Generación de interfaz básica para el software
 - a. Definición de la plataforma adecuada para la implementación.
 - b. Desarrollo de las herramientas en la plataforma seleccionada.
 6. Validación en contexto educacional del software básico para su uso.
 - a. Prueba con profesores
 - b. Prueba con Estudiantes en clase
 - c. Retroalimentación y cambios de aspectos analizados.

3.1. Obtención de datos.

Los datos que se utilizarán son obtenidos principalmente de libros e Internet escritos en Mapuzugun o que al menos contengan un mínimo de texto en esta lengua. Además, se usarán datos puestos a disposición en línea por investigadores de la Carnegie Mellon University, los cuales cuentan con datos transcritos y anotados de alrededor de 170 horas de relatos de Mapuzugun en distintas variantes territoriales.

Por otro lado, si fuese necesario, se pueden generar más datos a partir de textos y variaciones de ellos, como puede ser, cambio de grafemarios o data augmentation como lo visto en [18].

El preprocesamiento es principalmente separar las palabras, para probar tanto el detector de grafemarios como el segmentador, estas palabras vienen en distintas formas de escritura y podrían tener cambios en ciertos sufijos. Como ejemplo, en la Figura 3.1, para el equivalente al castellano de “la/lo dejaron vaciada” (una olla por ejemplo) se pueden ver las siguientes palabras que son equivalentes entre sí, pero tiene diferencias en su escritura (fondo rosado y celeste) y en el morfema kūno/kūnu dependiendo de la palabra, con fondo morado, siendo una diferencia fonética.

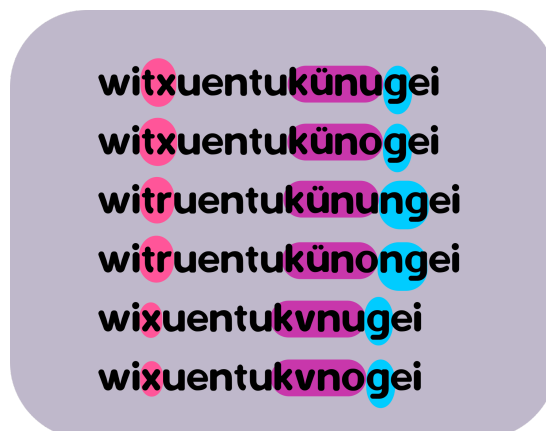


Figura 3.1. Diferencias posibles entre fuentes y grafemarios (Elaboración propia.)

Al encontrar alguna de estas palabras, su preprocesamiento es separar la palabra de su contexto para tomarla aislada. Este tipo de palabras sirve tanto para probar el detector de grafemario, si es que se tiene un origen claro y, además, para el segmentador si es que viene acompañado de una traducción.

También se necesitan textos completos con su traducción para poder hacer pruebas, al analizador morfológico y el traductor informal los cuales se preprocesan quitando las palabras en castellano, debido a que el programa no está pensado para entregar definiciones o análisis de palabras en ese idioma.

3.2 Traductor de grafemarios.

Esta primera herramienta tendrá la capacidad de recibir un texto cualquiera, escrito en uno de los grafemarios con más uso nombrados en este trabajo, para traducirlo a los otros sistemas de escritura. Además, que en el caso de ingresar un grafemario no definido como “más usado”, pueda hacer una extrapolación y generar, con un buen porcentaje de logro, una buena traducción de estos a alguno de los “conocidos”.

La herramienta funciona con base en un mapeo de los fonemas de los grafemarios y la diferencia entre estos. Para detectar un grafemario se debe traducir un texto desde un sistema de escritura hacia otro y luego volver al primero. El que no hayan diferencias entre el texto original y el que resulta luego de la segunda conversión indica que, o es del grafemario que se indicó en el comienzo, o que utiliza fonemas que son comunes para 2 o 3 grafemarios (dependiendo del texto), esto está en un diagrama en la Figura 3.2.

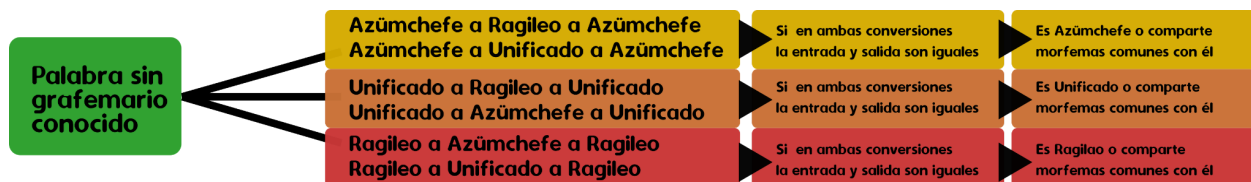


Figura 3.2. Diagrama del proceso de detección de grafemario (Elaboración propia.)

Caso 1: Detección de grafemario de una palabra en Azümcheffe, primero se asume Unificado y falla (Figura 3.3), luego se asume Azümcheffe y lo identifica (Figura 3.4).

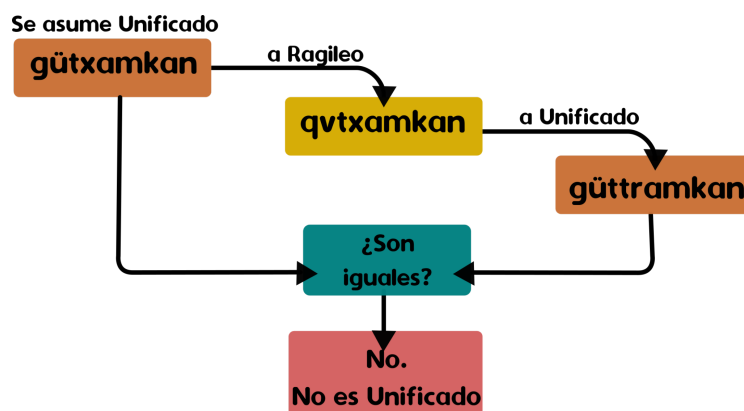


Figura 3.3. Caso Unificado equivocado (Elaboración propia.)

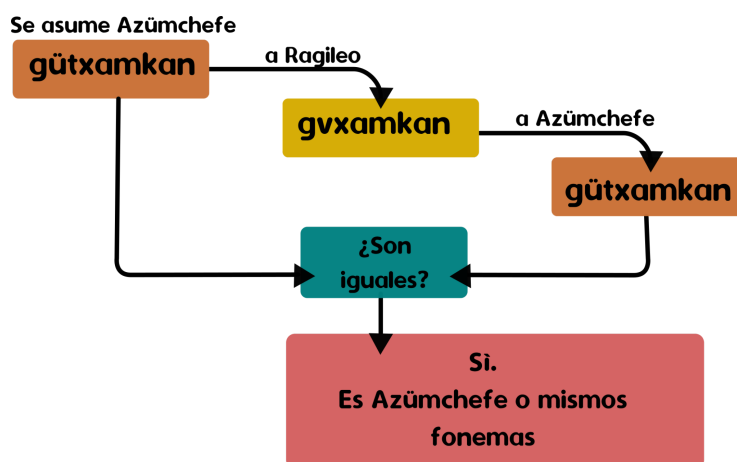


Figura 3.4. Caso Azümcheffe Correcto (Elaboración propia.)

Caso 2: Detección de una palabra escrita en grafemas comunes para los grafemarios (Figura 3.5).

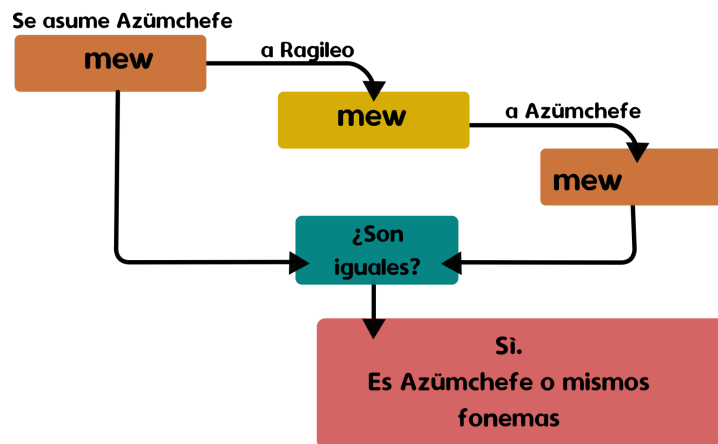


Figura 3.5. Caso Azümchefe coincidencia de fonemas (Elaboración propia.)

Al tener esta base y confirmar con dos variaciones cada vez, se puede detectar con una buena tasa de acierto el grafemario.

Además, se incluyeron las siguientes variantes que se asumen como los grafemarios “conocidos” pero con modificaciones en fonema. (Azümchefe, Unificado, Ragileo).

Debido a que se tienen reglas claras para poder hacer las conversiones entre grafemarios, se opta por hacerlo a través de expresiones regulares y no con inteligencia computacional, debido a que se tiene toda la información para convertir y, por lo tanto, detectar estos casos.

3.3 Analizador morfológico.

Como segundo paso, luego de la detección y normalización, se debe hacer una segmentación de la o las palabras que fueron ingresadas desde el traductor de grafemarios.

Para esto se tienen conjuntos de morfemas ordenados función: verbos, sustantivos, adjetivos, sufijos y terminaciones o desinencias, los cuales tienen un orden definido. En particular, para el caso de los sufijos, además, tienen un orden interno que está dado por las gramáticas consultadas en este trabajo.

Un ordenamiento general para una palabra en Mapuzugun sería el siguiente (Figura 3.6):



Figura 3.6. Ordenamiento general del mapuzugun. Palabra de ejemplo: kim kofkekefuygün (Elaboración propia.)

Un aspecto a considerar es que, si bien, estas descripciones son gramaticales y lingüísticas, se le intentará dar un enfoque comunicativo, por lo que se fusionarán ciertos sufijos para que sean entendibles por personas que no necesariamente tienen conocimientos técnicos sobre esta lengua.

Por ejemplo, algunas de las descripciones sobre transiciones incluyen al morfema “e” como separado de los morfemas que indican las personas gramaticales que participan en las transiciones. En el caso de este trabajo, se fusionarán la e con los distintos marcadores de personas para generar y presentar una unidad que aporte un sentido comunicativo.

Como ejemplo de lo anterior se tiene la palabra “eluen”, que significa “Tú me diste a mi algo”, para las separaciones gramaticales más formales tendríamos la separación elu-e-n, pero en el caso de esta herramienta se tiene la separación elu-en, en donde el morfema “en” representa una transición en modo indicativo de segunda persona singular a primera persona singular, esto interpretado se podría definir, para un contexto educativo del mapuzugun, como “una acción de tu a mi: ‘tú me ...’”.

El método para construir estas posibilidades será trabajar con árboles como el de la palabra AMULEN que se puede ver en la Figura 3.7. En este árbol, se puede ver que ambas palabras finales “AMU-LE-N” y “AMU-L-EN”, tienen un nodo común que es AMU, y de esta raíz verbal se desprenden las dos posibles segmentaciones.

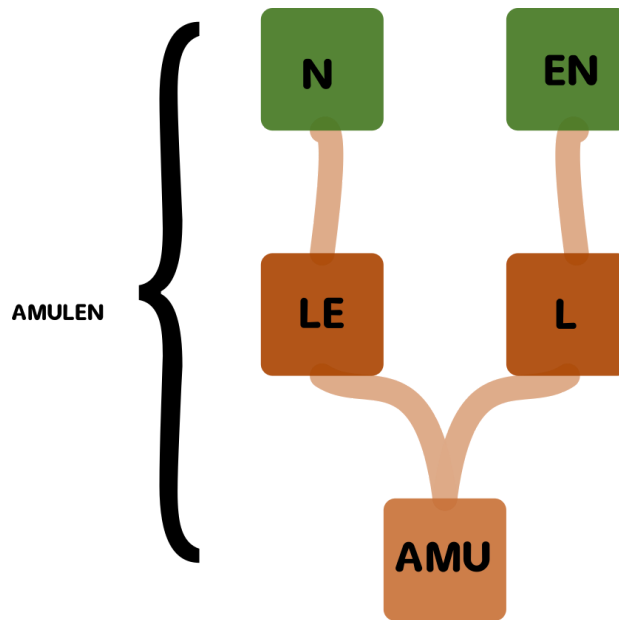


Figura 3.7. Muestra del árbol construido para la palabra amulen. (Elaboración propia.)

3.4 Traductor informal.

A partir de la separación, el “traductor informal”, presenta una traducción a los morfemas o conjuntos de estos para tener una noción de la separación morfológica que se está mostrando. Se define como informal debido a que el resultado de esta es un conjunto de definiciones al castellano que busca contener información suficiente para su inferencia por parte de una persona que no necesariamente esté completamente familiarizada con la lengua.

Al igual que en el analizador morfológico, en el que la separación en morfemas tiene un enfoque más comunicativo que técnico, las traducciones o significados de estas unidades también serán enfocadas en el uso más que en la lingüística. Este enfoque se puede ver reflejado en el sufijo -nie-, que Ineke Smeet define, en uno de sus usos, como “progresivo persistente” pero que a partir de su traducción como raíz verbal y de su uso en contextos de comunicación, se puede definir como “tener” o “estar” haciendo algo, dependiendo del verbo en el que vaya.

En un ejemplo aplicado del sufijo -nie- para la frase “azkintunien kiñe üñüm” que significa “estoy viendo/observando un pájaro” se puede ver en la Figura 3.8.

azkintu-	ver/observar
-nie-	progresivo persistente (L) tener / estar (C)
-n	primera persona singular (yo)
kiñe	un, uno, una
ũñũm	pájaro

Figura 3.8. Traducción de unidades de la frase “azkintunien kiñe ũñũm” (Elaboración propia.)

A partir de las definiciones mostradas para el sufijo -nie-, se decide optar por el tener, tomando en cuenta que si está presente en la traducción, y que, además, el término “progresivo persistente” no es de uso cotidiano para la mayoría de las personas que no tienen estudios de lingüística.

3.5 Integración de las herramientas.

Luego de generarse los módulos, se integran en un solo proceso (Figura 3.9):

1. Entra una palabra o frase.
2. Se detecta el grafemario a partir de las conversiones entre estos con el traductor de grafemarios.
3. Se tiene como salida, la palabra original y la palabra en Ragileo, además del grafemario detectado.
4. A partir de las salidas del detector, se toma la palabra en Ragileo y se introduce en el analizador morfológico.
5. Se genera como salida el árbol de posibles separaciones, con cada una de las palabras y sufijos se convertidas al grafemario original de entrada en el paso del detector de grafemarios, además de que se entregan las expresiones regulares utilizadas por el programa.
6. Para cada morfema detectado se le asigna una traducción o explicación al castellano.
7. Como salida del traductor informal y de la herramienta, se entregan todas las separaciones válidas en el grafemario original, junto con sus expresiones regulares y la asignación de traducciones o explicaciones.



Figura 3.9. Diagrama de integración de las herramientas. (Elaboración propia.)

3.6 Interfaz de usuario.

Para esta interfaz y para hacer el primer prototipo de este software se decide hacer una página web en el framework FLASK (Python), debido a lo compacto que pueden ser los códigos y a lo fácil su integración debido a que el desarrollo previo y el hecho durante la tesis está escrito en Python.

Se decide la página web como un intento de llegar tanto a dispositivos móviles como a navegadores web. Como se puede ver en la Figura 3.10 la interfaz incluye una explicación de la herramienta para contextualizar al usuario, luego el campo de uso de la herramienta, también se incluyeron ejemplos de textos y ejemplos de salidas del programa para que se puedan analizar en caso de que un usuario esté explorando sin tener palabras en mapuzugun para probar.

Además, se tiene una caja de comentarios la cual se habilitó para recibir retroalimentación tanto durante como luego del desarrollo de este trabajo.

Título Castellano	Título Mapuzugun
Saludo en Castellano	Saludo en Mapuzugun
Ingreso de palabras/frases INPUT Boton	
-Resultado desplegable 1 -Resultado desplegable 2 -Resultado desplegable 3 ... -Resultado desplegable N	
-Ejemplo de Resultado desplegable 1 -Ejemplo de Resultado desplegable 2 -Ejemplo de Resultado desplegable 3 ... -Ejemplo de Resultado desplegable N	
Texto de ejemplo en castellano para probar el software (para usuarios que no están familiarizados con la lengua)	Texto de ejemplo en Mapuzugun para probar el software (para usuarios que no están familiarizados con la lengua)

Figura 3.10. Bosquejo de la interfaz del software. (Elaboración propia.)

3.7 Prueba con usuarios.

La prueba con usuarios consiste en una validación del software que se divide en general en 3 partes: Cuánto apoya al aprendizaje, la usabilidad de éste, y los comentarios generales que se puedan obtener de las pruebas.

Desde el punto de vista del aprendizaje se evaluarán traducciones de palabras o textos más largos. Para observar si es que existe una mejora en las traducciones, a partir de una comparación del desempeño sin el uso de la herramienta en contraste con el uso de esta.

Sobre la usabilidad se pretende obtener datos de la facilidad de uso del software además de evaluaciones generales y sobre la parte visual de este.

Finalmente, los comentarios generales tienen la particularidad de que no están estandarizados y por lo tanto pueden aportar datos y/o visiones del software que no necesariamente están contenidos dentro de las pruebas.

Uno de los posibles obstáculos para esta parte tiene que ver con la cantidad reducida de personas que podrían estar interesadas o dispuestas, esto considerando que la cantidad de hablantes y personas relacionadas a la lengua podría ser pequeña comparada con la población del país. Pero además, podrían haber personas que desconfían de ciertas instituciones al momento de trabajar con sus propias prácticas culturales, en este caso, la lengua.

4. Traductor de grafemarios.

En esta sección se habla de como está concebido y construido el Traductor de grafemarios construido, este consta de dos partes: Las reglas de paso de un grafemario a otro y el algoritmo de detección que se usa para poder hacer una predicción de grafemario.

4.1 Reglas de paso de grafemarios.

Para comprender las reglas con las que se pasa de un grafemario a otro, lo primero que se debe entender son las diferencias y coincidencias entre los grafemarios. Los tres grafemarios en los que se enfoca este trabajo fueron escogidos debido a que son los más usados. Estos son:

Grafemario Ragileo de 1985 creado por Anselmo Ragileo,
Alfabeto Mapuche Unificado de 1989 creado por Maria Catrileo y
Grafemario Azümchefe creado por Necul Painemal para la CONADI.

Entre estos tres grafemarios se tienen 16 coincidencias, definidas para este caso como sonidos y letras coincidentes, la siguiente lista muestra los grafemas mencionados:

A, E, F, I, K, L, M, N, Ñ, O, P, R, T, U, W, Y.

Estas letras son coincidentes independientemente de si están escritas con mayúsculas o minúsculas. Otra cosa importante que se debe hacer notar sobre las coincidencias es que, en términos de la detección, no aportan información sobre el grafemario en el cual está escrita la palabra, en tanto estas reglas de conversión se centran en las diferencias de grafemarios.

4.1.1 Diferencias de grafemarios.

Para describir las diferencias de escritura, al ser tres grafemarios, existen seis posibles reglas para ir de cada grafemario a alguno de los otros dos como indica la siguiente Figura 4.1:

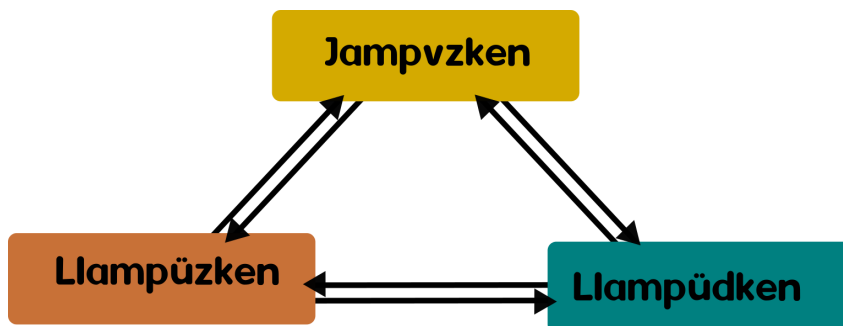


Figura 4.1. Pasos entre grafemarios (Elaboración propia.)

En esta Figura, “*Jampvzken*” está escrito en Ragileo, “*Llampüdken*” en Unificado y “*Llampüzken*” en Azümchefe, las tres refiriéndose a la misma fonética en Mapudungun de la

palabra “Mariposa” del castellano. En este ejemplo se puede notar la relación entre la J en Ragileo con la Ll en Azümcheffe y Unificado.

En total, para la relación Unificado-Ragileo, se tienen 10 diferencias que se muestran en la siguiente Tabla 4.1, en la siguiente caso Unificado-Azümcheffe se tienen 8 diferencias, Tabla 4.2 y en la relación Ragileo-Azümcheffe se tienen 8 diferencias Tabla 4.3.

Unificado	Ragileo
CH	C
D	Z
G	Q
<u>L</u>	B
Ll	J
<u>N</u>	H
Ng	G
Tr	X
<u>T</u>	-
Ü	V

Tabla 4.1. Diferencias entre grafemarios Unificado y Ragileo (Elaboración propia.)

Unificado	Azümcheffe
D	Z
G	Q
<u>L</u>	Lh
<u>N</u>	Nh
Ng	G
Tr	Tx
<u>T</u>	T'
S	Sh

Tabla 4.2. Diferencias entre grafemarios Unificado y Azümcheffe (Elaboración propia.).

Ragileo	Azümcheffe
<u>C</u>	Ch
<u>B</u>	Lh
J	Ll
H	Nh
S	Sh
X	Tx
-	T'
V	Ü

Tabla 4.3. Diferencias entre grafemarios Ragileo y Azümcheffe. (Elaboración propia.)

4.1.2 Orden de conversión.

A partir de estas diferencias, se debe tomar especial precaución en los grafemas que tienen otro grafema como subcadena. Un ejemplo de esto es en el grafemario Unificado el grafema “Ng”, que además contiene el grafema “G” que se usa mismo sistema de escritura para otro fonema. O, casos en los que tres grafemas contienen una misma letra, que además es un grafema por separado, como la letra “L” en “L”, “Lh” y “LL”. El único grafemario que no tienen este problema interno es Ragileo, y esto sucede porque el autor usa una letra para cada fonema del Mapuzugun. Esto permite pasar más rápidamente de Ragileo a algún otro grafemario, siempre cuidando el orden de las transformaciones cuya salida puede generar una ambigüedad de morfemas durante la conversión.

El orden que se debe tener en cuenta tiene que ver con que si un morfema es contenido por otro, primero se debe desambiguar ese caso y luego seguir con los demás cambios.

En el caso de la Ng y g, para ir de Unificado a Ragileo o Azümcheffe, siempre que haya una g y no haya una N precediéndole, se puede cambiar a Q, por lo tanto antes de hacer las transformaciones se debe probar la Ng, guardando el resultado (“G”) de esta en una variable auxiliar para poder convertir luego todas las G del Unificado a Q. Hecho este último paso se quita la variable auxiliar y se vuelve a poner la G resultante del cambio.

Como ejemplo se deja un trozo de código que tiene el procedimiento de Unificado a Ragileo para que se pueda ver como funciona.

```
txt=re.sub('ng','d',txt)    #Ng,ng    G,g
txt=re.sub('g','q',txt)    #G,g    Q,q
txt=re.sub('d','g',txt)    #D,d    G,g
```

Profundizando en esto, se puede ver que siguiendo este orden la expresión “se volvió blanco”, que en Unificado es lig ngetuy , se puede pasar a Azümchehe siguiendo esas reglas:

1. Palabra “lig ngetuy”
2. paso “ng” a d => “lig detuy”
3. paso g a q => “liq detuy”
4. paso d a g => “liq getuy”

Para esta misma frase, si no utilizamos la variable “d”:

1. Palabra “lig ngetuy”
2. paso “ng” a g => “lig getuy”
3. paso g a q => “liq qetuy”

O invirtiendo el orden de g y ng en sus respectivos pasos:

1. Palabra “lig ngetuy”
2. paso g a q => “liq ngetuy”
3. paso ng a g => “liq ngetuy”

En los anexos se adjunta el código correspondiente a los seis pasos entre grafemarios.

4.2 Detector.

Para hacer el detector se utilizan las reglas que se construyeron a partir de diferencias entre los grafemarios. A partir de las reglas, se harán conversiones en dos sentidos para comprobar si es que la palabra es candidata para estar escrita en alguno de los grafemarios que estamos trabajando en este documento.

El algoritmo es el siguiente para la comprobación de un grafemario **G1**, probando con la palabra **w** y los grafemarios auxiliares **G2** y **G3**. Este algoritmo se denominará **A1**:

1. Se asume que **w** está escrita en el grafemario **G1**.
2. Se convierte **w** al grafemario **G2** a partir de las reglas, resultando **w'**.
3. Se convierte **w'** al grafemario **G1**, resultando **w''**.
 - a. Si **w** $\neq$ **w''**:
La palabra no está escrita en **G1**, se descarta este grafemario como posible.
 - b. Si **w** = **w''**:
La palabra podría estar escrita en el grafemario **G1**.
Se continúa.
4. Se convierte **w** al grafemario **G3** a partir de las reglas, resultando **w'**.
5. Se convierte **w'** al grafemario **G1**, resultando **w''**.
 - a. Si **w** $\neq$ **w''**.
 - i. La palabra no está escrita en **G1**.
 - b. Si **w** = **w''**:
 - i. Se asume que **w** está escrito en **G1** (no excluyente de estarlo en **G2** y **G3**).

El proceso descrito se hace para cada grafemario, pudiendo resultar en la detección de 0, 1, 2 y 3 grafemarios posibles para la palabra analizada.

Como ejemplo de 0 detecciones, está la palabra “*chacay*” que entregaría un resultado negativo para los 3 grafemarios estudiados, debido a que la *c* no aparece sola ni en Unificado ni en Azümcheffe y solo aparece en Ragileo, sin embargo el grafema *ch* no existe en este último grafemario. A partir de este resultado, no se puede concluir nada sobre el grafemario en que está escrita esta palabra.

Para describir un ejemplo de una detección de grafemarios, se pueden mostrar las palabras “*Llampüzken*”, “*Jampvzken*” o “*Llampüdken*”, que están escritas, únicamente, en Azümcheffe, Ragileo y Unificado respectivamente. Cada una de estas palabras tiene solo un posible grafemario a partir de este detector.

En el caso de la detección de 2 grafemarios, se pone como ejemplo la palabra *mawün*, que debido a que tiene el grafema “*ü*”, descarta a Ragileo y los grafemas restantes son los coincidentes para los tres grafemarios. Debido a esto, existe una ambigüedad que se podría resolver al ingresar palabras de contexto.

Finalmente para los casos en que solo existen grafemas coincidentes de los 3 grafemarios, se pueden detectar los 3 grafemarios, como es el caso de las palabras “*amun*” y “*mew*”.

Para tratar de validar este detector, se tomaron varios textos, se separaron en palabras y se clasificaron de acuerdo a su grafemario original (Ragileo, Azümcheffe o Unificado, dependiendo del texto), se quitaron palabras en castellano, nombres de personas y siglas, que pueden provocar ruido en esta clasificación, también se eliminan duplicados que tuvieran escritura y clasificación real iguales. Por otro lado, se quitaron todas las palabras que tienen solo las letras coincidentes de los 3 grafemarios al mismo tiempo. Se dejaron, entonces, sólo las palabras que, a lo más, podrían existir en 2 de los 3 grafemarios. En total la cantidad de palabras que se procesaron para ver esta prueba fueron 24.599 de los cuales un 22.973, correspondiente a un 93.4%, fueron clasificadas de forma correcta según el grafemario de origen de su texto.

4.2.1 Posibles variaciones.

Para los casos sin detectar, se trabajó en una forma de ampliar el espectro de búsqueda de los grafemarios. A partir de los 3 grafemarios que ya se tienen, se pueden aplicar ciertos cambios antes de hacer el algoritmo de detección para aumentar las posibilidades de entender la escritura que se usó en una palabra o texto.

La forma de hacer esto es pre-procesar las palabras, de tal forma que se modifica un posible grafema y se hace la prueba de detección:

Por ejemplo la palabra “saludarse” podría ser encontrada en textos como “*chaliwvn*”, que no está escrita en siguiendo las reglas de ninguno de los 3 grafemarios. Una forma de ver esta palabra es entenderla como una palabra escrita en Ragileo, al cual se le reemplazó el grafema C con el CH y, por lo tanto, si es que para esta palabra se transforma CH a C, calzaría con las reglas de Ragileo. Esto significa que es una variante de Ragileo.

Esto se puede expresar en un algoritmo (**A2**) que utiliza además el algoritmo **A1** (detección de grafemarios), para la palabra **w**, el grafemario que, se asume, tiene la palabra **G1**, los grafemarios **G2** y **G3** que son los denominados grafemarios auxiliares, y la variación de un grafema de **G1**, llamado **V1**:

1. Ingreso la palabra **w** en el algoritmo **A1**.
 - a. Si tiene resultados:
Termina el proceso
 - b. Si no tiene resultados:
Se continua
2. Se aplica la variación del grafema **V1** a **w**, resultando en **w'**.
3. Se aplica el algoritmo **A1** a la palabra **w'**.
 - a. Si no tiene resultados:
La palabra no está escrita en **G1** con la variación **V1**. Y termina el proceso.
 - b. Si tiene resultado positivo para **G1**.
Se revierte la variación **V1** para **w'**, resultando **w''**.
 - i. Si **w ≠ w''**:
La palabra **w** no está escrita en **G1** con la variación **V1**. Termina el proceso.
 - ii. Si **w = w''**:
La palabra podría estar escrita en **G1** con la variación **V1**.

Este algoritmo **A2**, se prueba para cada combinación de grafemario (**Gn**) y sus variaciones correspondientes (**V1** a **Vn**)

Las variaciones que se incluyeron en este estudio fueron:

1. Ragileo con Ch reemplazando a C.
2. Ragileo con Ñ reemplazando a V.
3. Ragileo con D reemplazando a Z.
4. Unificado con Z reemplazando a D.
5. Unificado con Tx reemplazando a Tr.

Estos se eligieron para hacer la prueba de variantes. Sin quitar la posibilidad que estas mismas variantes puedan tener su propia denominación como un grafemario aparte.

Hecha la detección del grafemario de la palabra, se puede hacer una normalización de grafemarios desde cualquiera de los 3 principales o sus variaciones a cualquiera de los 3 principales o sus variaciones. Esto permite que los procesos que se puedan hacer después no se ocupen del grafemario, sino que se ocupen de su tarea específica, como en el caso del analizador morfológico.

Además, esto puede ser de utilidad para una persona que no esté familiarizado con los múltiples sistemas de escritura, evitando la confusión de dos palabras que se escriben distinto pero que podrían ser en realidad, fonética y semánticamente hablando, la misma palabra.

5 Analizador morfológico.

Este analizador morfológico separa las palabras y las presenta como composición de morfemas. Es una parte fundamental de la herramienta. A continuación se muestra su proceso.

5.1 Desarrollo.

Para generar este analizador, se deben separar palabras y, para esto, se tienen claramente definidas las estructuras del Mapuzugun, a nivel lingüístico en las gramáticas que se han escrito [35,39,57], por lo tanto se tomaron y se implementaron a través de expresiones regulares, como otra fuente se tomó la recopilación que se hizo en azumchefe.cl de la gramática del idioma [59].

Se trabajó con conjuntos de expresiones regulares que contienen los morfemas del idioma. Estos conjuntos separan, por función: en raíz verbal, sustantivo/adverbio/adjetivo, sufijos, y desinencias. Además, la posición juega un papel importante, porque cada uno de los morfemas tiene una posición particular a la que [35] llama “slots” en inglés, esto se puede llevar al castellano como casilleros.

A partir de estas expresiones regulares se recorre la cadena de una palabra y se genera su árbol. Cada rama se evalúa para ser considerada como válida, y por lo tanto pasar al siguiente paso de “traductor informal”. Los morfemas y su orden deben cumplir ciertas restricciones que tienen que ver con la correcta formulación de palabras en mapuzugun, tanto en orden, como se mencionó antes, pero también en la compatibilidad de que dos morfemas estén en la misma palabra.

Para este módulo, se escogió el grafemario Ragileo para hacer la separación, por lo tanto cualquier palabra de otro grafemario que quiera pasar por el analizador debe, necesariamente, pasar por el conversor de grafemario. Esta decisión tiene que ver con que, como vimos antes, cada sonido del Mapuzugun se puede asociar a una sola letra en particular, por lo que es más sencillo pasar de Ragileo a cualquier otro grafemario.

El algoritmo para la generación del árbol mencionado es el siguiente:

5.2 Algoritmo.

0. Pre-procesamiento de la palabra, se convierte minúsculas y , en caso de que sea necesario, se escribe en Ragileo para su introducción en el programa.
1. Se recorren las expresiones regulares para sustantivos/adverbios/adjetivos y raíces verbales:
 - a. Si la expresión calza con una cadena completa o una subcadena, se guarda como rama para este nivel.
2. Se evalúan las ramas:
 - a. si no existen ramas, se termina el proceso sin resultado.
 - b. si existen ramas se continúa.
3. Para cada rama:
 - a. Se quita la subcadena correspondiente a la expresión regular con la que se detectó la rama, y se analiza el resto de la cadena.
 - i. Si no hay cadena:
 1. Si la última expresión regular es del grupo “desinencias” la rama se da por finalizada y se continúa con la siguiente.
 2. Si la última expresión regular no es del grupo “desinencias” la rama no es válida y se elimina. Se continúa con la siguiente rama.
 - ii. Si hay cadena se continúa el proceso.
 - b. Se recorren las expresiones regulares para sustantivos/adverbios/adjetivos, raíces verbales, sufijos y desinencias.
 - i. Si la posición de orden es menor que la de la raíz de esta rama:
 1. No se prueba y se continúa con la siguiente.
 - ii. Si la posición de orden es mayor o igual que la de la raíz de esta rama:
 1. Si la expresión calza con la cadena completa o una subcadena, se guarda como subrama para este nivel.
 - c. Se evalúan las ramas:
 - i. si no existen subramas, termina el proceso sin resultado y se elimina la rama padre.
 - ii. si existen ramas se continúa.
 - d. Para cada subrama se aplica el paso 3, hasta que no quedan posibles combinaciones de expresiones regulares.
4. Se retorna el árbol.

Se debe también aclarar que los morfemas también se pueden adaptar para una mejor comprensión por los usuarios, en este caso en algunas publicaciones la letra -l- se considera un sufijo cuando se ve como en la palabra “amu-l-mi” que significa “si (tú) vas”, pero para efectos del analizador, que tiene un enfoque comunicativo, se considera -lmi como una desinencia completa incluyendo también a la -l-, porque implica en la traducción que es una acción condicional que hace la segunda persona singular (tú) amu-lmi. La separación de este ejemplo se modifica por la forma en la que se enseña el mapuzugun y no en la forma en la que se analiza técnicamente.

Al analizador morfológico pueden ingresar palabras simples o con sufijos y se pueden identificar sus segmentos, además, con anotaciones de los segmentos, se deberá entregar información de raíces y/o sufijos que contiene la palabra. Bajo la misma lógica se pueden ingresar frases, las cuales serán analizadas por palabra, pero mostradas en una misma respuesta sin necesidad de hacer una nueva consulta para cada palabra que se tenga en la frase.

La idea inicial era adaptar un analizador morfológico ya desarrollado por Andrés Chandía y disponible a través de internet. De este, se tomaría su salida y se simplificaría, pero al desarrollar este trabajo y para facilitar la integración con el detector de grafemarios, se decidió que se tomarían fuentes comunes [35] con este analizador y se desarrollaría uno nuevo que tenga en consideración los factores de educación (el cómo se enseña) y de comunicación (el cómo se utiliza) complementando con otras fuentes [39,57] como la gramática ineke.

5.3 Traductor Informal.

Luego de hecha la segmentación se implementó la parte del traductor informal, que asigna definiciones a los morfemas y, en algunos casos, a combinaciones de morfemas dependiendo de su función comunicativa. Para esto se tomaron todos los morfemas y se les asignó una definición tratando de ser lo más claro posible y dejando fuera en lo posible las definiciones técnicas complejas. Además, para el caso de sustantivos, verbos y adjetivos se dejó la definición que tienen en un diccionario llamado “Ziccionario Mapuzugun – Castellano”, el cual está disponible en la web [58].

Como ejemplo de una traducción tenemos la separación de la palabra “txekayawkelai”, en la que una de las posibles separaciones sería:

txeka-yaw-ke-la-i

Según esta forma de dividir, los componentes de la palabra son:

txekan-:	vi & vtr caminar, marchar, pasear , vtr medir con pasos
-yaw-:	andar
-ke-:	habitualmente
-la-:	negación modo "normal"(indicativo)
-i:	El/Ella

A partir de esta división de la palabra “txekayawkelai”, se espera que una persona llegue a la frase “el/ella no anda caminando habitualmente”. Sobre esto último es que se trata probar que tan efectivo puede llegar a ser. La prueba verá, si es que con la información entregada en la explicación o traducción informal, un usuario se puede acercar o puede acertar la traducción correcta para las palabras que se introduzcan en el software.

Los desafíos de esta parte van, principalmente, en que la experiencia de la persona con respecto al Mapuzugun, podría incidir en su interpretación de las traducciones, por lo que deben ser lo más claras posibles para que esa dependencia no sea tan fuerte. También se deben tomar en cuenta ciertos casos de partículas que si bien actúan en conjunto, podrían estar separadas dentro de la palabra. Como ejemplo de morfemas que actúan en conjunto se tienen, para las transiciones de segundas personas a primeras personas, exceptuando entre singulares, en este caso el morfema MU indica que existe una transición, y las terminaciones “n”, “yu” y “iñ” nos indican sobre quién es la acción, aquí se pueden ver ejemplos:

elumun kofke (elu-mu-n kofke): ustedes me dieron pan

elumuyu kofke (elu-mu-yu kofke): ustedes nos dieron pan a nosotros 2 o
tú nos diste pan a nosotros 2

elumuiñ kofke (elu-mu-iñ kofke): ustedes nos dieron pan a nosotros (más de 2) o
tú nos diste pan a nosotros (más de 2)

Este caso es sencillo e incluso podrían tomarse como morfemas de transición “muiñ”, “muyu”, “mun”, pero estos morfemas no siempre están juntos, como en el caso de:

kofke elumun :	ustedes me dieron pan
kofke elumuan :	ustedes me darán pan
kofke elumulan :	ustedes no me dieron pan
kofke elumulayan :	ustedes no me darán pan
kofke elumulayafun :	ustedes no me darían pan
kofke elumurpulayafun :	ustedes no me pasarían a dar pan

En este caso los morfemas “muiñ”, “muyu” y “mun” pierden sentido.

Aunque, en casos como “pe” y “la”, que separados hacen un pasado y una negación respectivamente, “amup**elai**” (amu-pe-la-i) podría significar “no ha ido” o “no fue”, pero también podría significar (amu-pela-i) “debe haber ido”. En este caso sí podría generarse un morfema auxiliar que aglutina ambos morfemas, debido a que no aparecen separados.

6. Implementación del software.

Para la construcción de este software, se integraron los módulos descritos en los capítulos anteriores de forma que se procesa cada palabra pasando sucesivamente por cada herramienta.

Como input de parte del usuario se tiene una palabra sin procesar, esta palabra pasa a la detección de grafemario, luego al análisis morfológico y, finalmente, se traducen los morfemas, generando una palabra segmentada que muestra las traducciones de sus morfemas.

La integración fue en un servidor web provisorio que sirvió para montar el software y poder mostrarlo a las personas que participaron de las pruebas con usuarios. El software fue desarrollado en el framework Flask para Python, este fue escogido por su simplicidad para hacer aplicaciones web pequeñas que no necesiten una infraestructura importante.

Al ser un framework para Python, se facilita la integración con los códigos de las herramientas, las cuales fueron escritas en ese mismo lenguaje.

6.1 Proceso de una palabra.

El proceso de una palabra parte en un cuadro de texto para ingresar las palabras al software marcado en rojo en la Figura 6.1, al hacer clic en botón, se inicia el la secuencia de los módulos. Esta palabra se hace minúscula y se quitan símbolos que puedan aparecer para su correcto procesamiento.

La palabra normalizada entra al detector de grafemarios el cual retorna el sistema de escritura en el que podría estar escrita la palabra. Además, independiente del grafemario detectado, se transforma a Ragileo, el cual fue escogido para el procesamiento del analizador morfológico. Se guarda en este paso el grafemario original y la palabra en Ragileo para el siguiente paso.

La palabra en Ragileo entra al Analizador morfológico y devuelve un árbol en Ragileo. Todos los componentes de Mapuzugun en los nodos (morfemas) se convierten al grafemario original. Estos morfemas llevan, también, una traducción de cada uno al castellano.

Luego esto se envía a la vista para que el usuario pueda ver la separación y la traducción de los segmentos de la palabra. Esto se puede ver en la Figura 6.2 para la palabra: “puwülgei”, escrita en Azümcheffe, también en la Figura 6.3 se puede ver que estas explicaciones se pueden cerrar, en caso de que el usuario no quiera ver las traducciones de los morfemas. Además, se pueden ver en la Figura 6.4 la misma palabra pero en Ragileo respectivamente. Independientemente del grafemario, estas palabras mostrarán la misma separación. Estos resultados se muestran en una pestaña dentro de la aplicación web, junto a ésta hay otras dos pestañas.

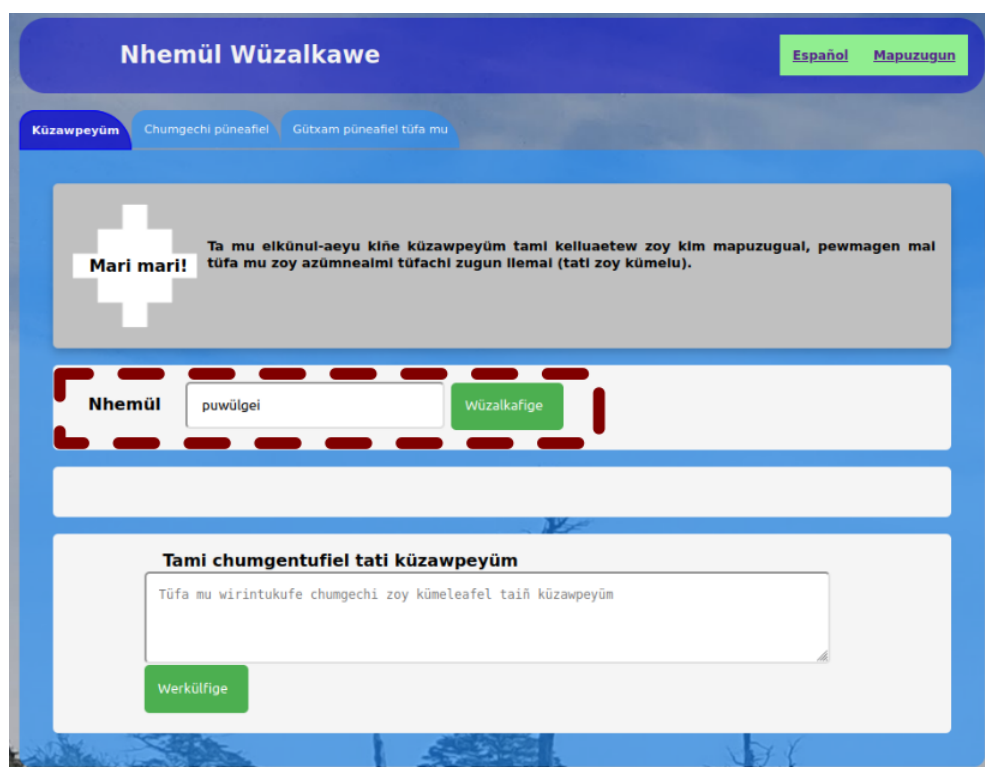


Figura 6.1 Entrada de una palabra al software (marcado en un cuadro rojo hecho de guiones).

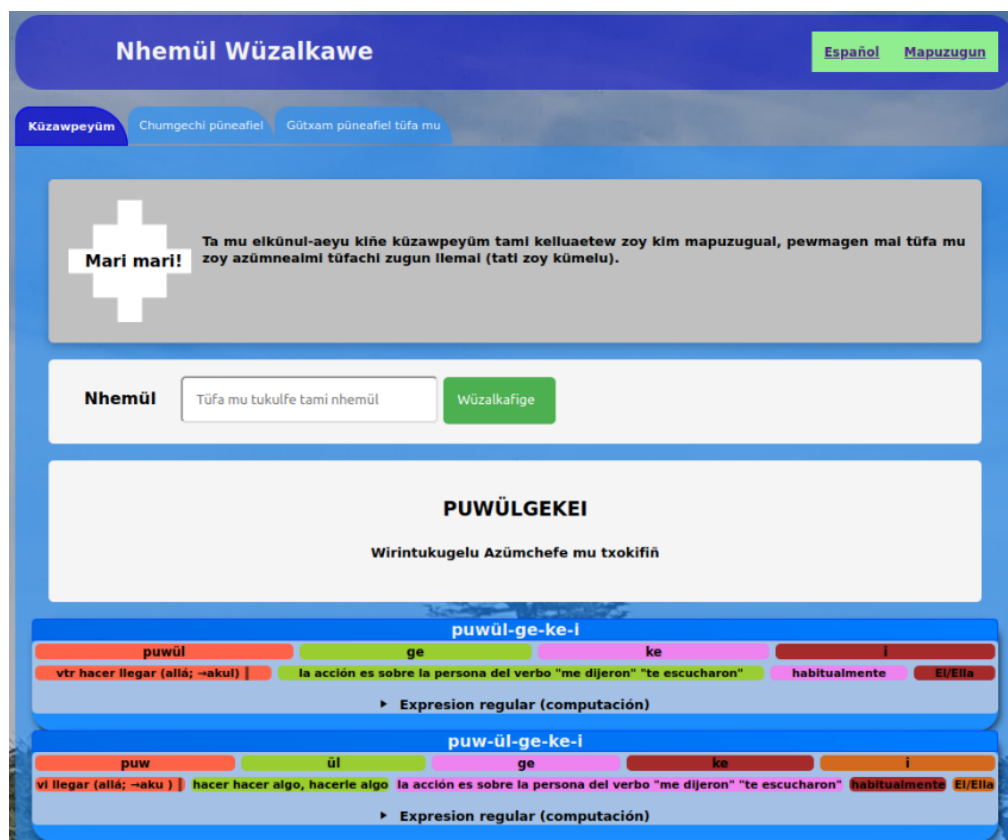


Figura 6.2. Muestra del resultado para la palabra “puwülgekei”, detectado como Azümchehe.



Figura 6.3. Resultado para la palabra “puwülgekei”, con una opción escondida.

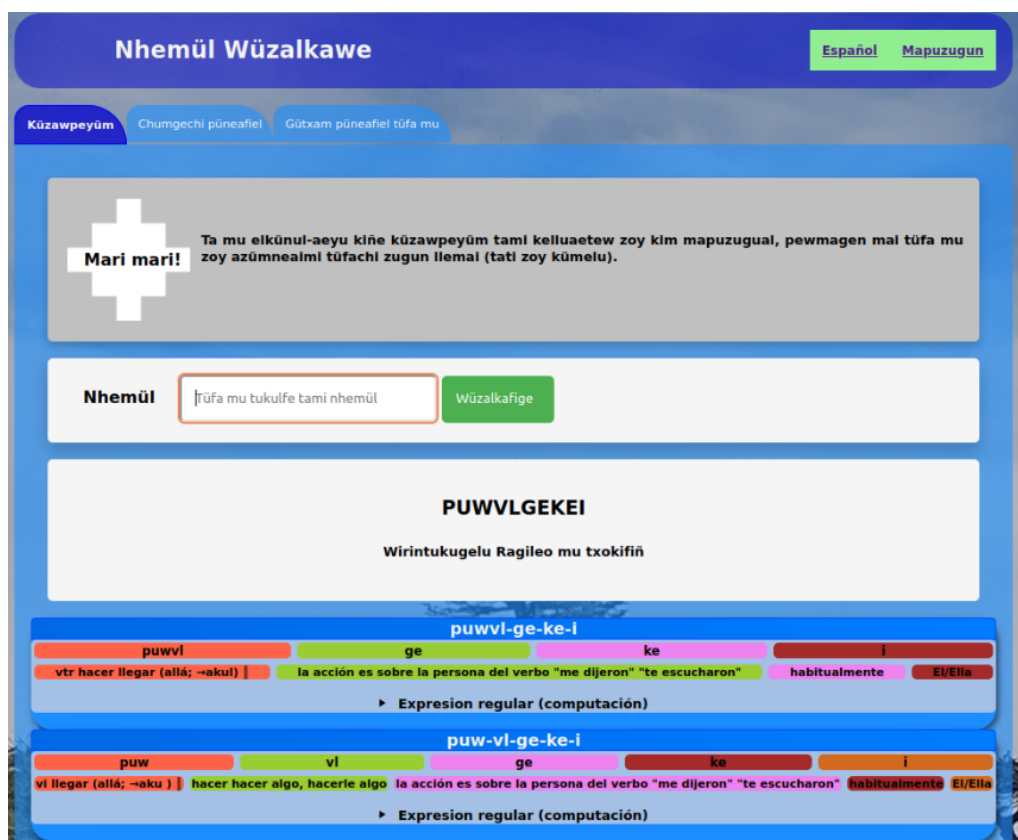


Figura 6.4. Resultado para la palabra “puwvlgekei”, detectado como Ragileo.
(Misma palabra que “puwülgekei”)

A partir de los resultados, mostrados en esta pestaña, un usuario tendrá información para intentar inferir cuál podría ser la traducción de esta palabra, en este caso “puwülgekei” significa “habitualmente se hace llegar allá”.

Para la palabra analizada se obtuvieron dos opciones pero es posible que internamente haya habido más potenciales resultados que no hayan podido cumplir las reglas gramaticales del Mapuzugun y, por lo tanto, fueron eliminadas como opciones válidas.

Otro punto importante, es que es posible cambiar el idioma de la interfaz para que se pueda utilizar el programa en castellano, para los usuarios que aún no tienen un nivel suficiente de Mapuzugun para entender el software en el idioma Mapuche, como se puede ver en la Figura 6.5.



Figura 6.5. Interfaz en español.

Al final de esta sección hay una caja de comentarios para poder tener una retroalimentación permanente y que ésta no dependa de encuestas o pruebas de usuario para hacer mejoras al software.

Luego, en la segunda pestaña se muestran un ejemplo de palabras segmentadas con la opción de mostrar su explicación y, más abajo, sus traducciones para guiar el uso de la herramienta, como se ve en la Figura 6.6. Finalmente, la tercera pestaña contiene un texto de ejemplo en Mapuzugun y luego en castellano, con el objetivo de que se puedan sacar palabras para utilizar en la herramienta. Estos textos se pueden ver en la Figura 6.7.

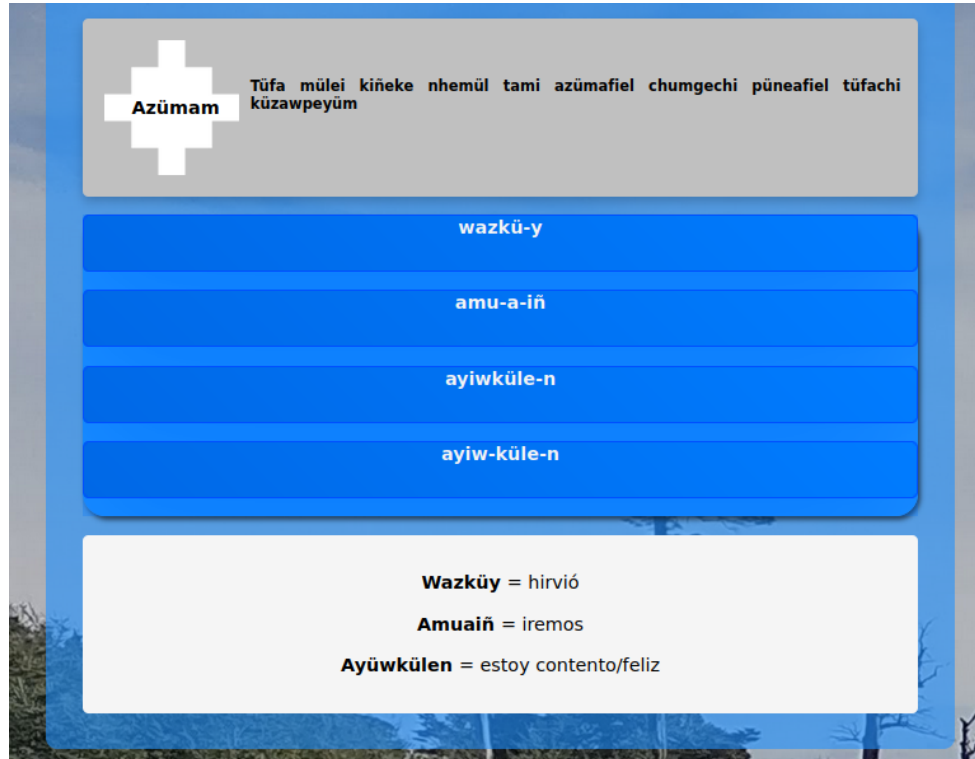


Figura 6.6. Página de ejemplos de uso de la herramienta de la interfaz actualizada.



Figura 6.7. Texto de ejemplo para probar la herramienta.

7. Pruebas del software

El software desarrollado se probó con diferentes grupos de usuarios. En general fue complejo encontrar estos grupos, debido a varias circunstancias. Entre las dos más importantes, mencionemos, primero, las características de los cursos actuales sobre Mapudungun. Hay muchos cursos que se dictan por diferentes organizaciones, pero en general sus contenidos, niveles y formas pedagógicas son muy diferentes, y existe poca sistematización de niveles. Esto dificultó el encontrar los grupos adecuados para probar el software. En segundo lugar, fue un problema la gran diversidad de formas de aprendizaje del Mapuzugun hoy día. Hay muchas personas que siguen cursos de manera intermitente, retoman cursos en otros lugares, o simplemente son autodidactas con materiales y personas de apoyo, y gran parte de esto en grupos y listas de Internet. Sumado a que, existen sectores que tienen una desconfianza sobre instituciones que históricamente han hecho lo que se denomina “extractivismo cultural”, que se puede ver reflejado en una gran cantidad de estudios desde la academia, que muchas veces quedan simplemente en el documento, pero que no tienen un buen manejo de expectativas o que no hacen una retribución a las comunidades desde donde se “extrae” la información. Estas dificultades propias de lenguas minorizadas ayudan a entender el diseño de las pruebas que se presentan en este capítulo.

Se realizaron dos pruebas que obtuvieron tres tipos de información:

En primer lugar, el apoyo lingüístico que se obtiene desde la herramienta, preguntando por traducciones de palabras, frases o textos, los cuales se intentaron traducir en una primera instancia sin la herramienta y luego para comparar se utilizará la herramienta. Las traducciones se evaluaron con notas de 0 a 1 y de 0 a 3 puntos, en la primera y segunda prueba respectivamente. En caso de notar un aumento de puntaje se toma como una indicación de que el software apoya en el aprendizaje.

Luego, se analiza la usabilidad del software a partir de preguntas de evaluación de la dificultad de uso y del aspecto visual de este.

Finalmente, los comentarios generales, que están presentes en todas las secciones, aportan información que no está contenida en la encuesta y, por lo tanto, información que puede ser relevante sobre todo para hacer las mejoras tanto a la interfaz como a la parte técnica-lingüística del software.

En una primera prueba con usuarios, se hizo una convocatoria a más de 200 personas que habían participado durante los últimos 5 años en cursos de mapuzugun. De las 200 personas se inscribieron 30 para probar la herramienta a través de una encuesta y de estas se obtuvieron 9 respuestas.

Dentro del grupo, hubo 3 personas con conocimiento avanzado del Mapuzugun, definido bajo el criterio que las tres personas, están enseñando o han enseñado la lengua en algún curso o taller.

Otra de las personas tiene un nivel intermedio, esto se puede ver reflejado en que en su puntuación se ve que responde más de la mitad de las palabras sin ayuda, aunque no tenga una puntuación perfecta. Además, existe una persona que no respondió ninguna de las preguntas. Los participantes que no fueron mencionados, tienen resultados variables, pero no son expertos ni principiantes completamente según la evaluación de sus traducciones.

Luego, la segunda versión se publicó en grupos de personas que están interesadas en Mapuzugun en redes sociales participando, esta vez, 32 personas de la encuesta. Las principales diferencias es que, en esta prueba, se hizo la pregunta respecto del nivel de Mapuzugun en el que se identifican los participantes y por lo tanto se tienen 5 grupos dentro de los cuales se clasifican estos usuarios:

0. "Estoy empezando a aprender. Todavía no sé mucho Mapuzugun."
1. "He estudiado un poco o he ido a cursos. Sé saludar y presentarme, pero no entiendo bien las conversaciones en Mapuzugun."
2. "He estudiado o he ido a cursos. Y puedo entender conversaciones de otras personas en Mapuzugun."
3. "He estudiado o he ido a cursos. Puedo mantener conversaciones en Mapuzugun."
4. "Soy hablante desde pequeñ@"

Para profundizar la descripción del segundo grupo, los 32 participantes se dividen de la siguiente forma:

- 6 personas en el grupo 0.
- 8 personas en el grupo 1.
- 7 personas en el grupo 2.
- 8 personas en el grupo 3.
- 3 personas en el grupo 4.

Otra diferencia importante es que se disminuyó la cantidad de preguntas para que la encuesta fuera menos tediosa desde el punto de vista del tiempo.

7.1 Encuesta

7.1.1 Primera prueba.

Parte 1:

Primero el usuario debe traducir nueve palabras, cada una de estas palabras tiene tres campos, estas palabras no están en un grafemario en particular si no que en los distintos mostrados en este trabajo, para comprobar su utilidad en casos en que hay textos o frases que vienen de distintas fuentes:

Palabra lhali (Traducirlo sin ayuda)

Texto de respuesta breve

Palabra lhali (Traducirlo apoyado en el programa)

Texto de respuesta breve

Escribe si tienes comentarios sobre la palabra y su separación en el programa (para la palabra "lhali")

Texto de respuesta breve

Figura 7.1. Ejemplo de preguntas para una palabra.

- **Campo 1 de palabra:**
Se debe traducir la palabra sin ayuda. Este campo indica cuánta información previa tiene el usuario sobre la palabra en cuestión para considerarlo al momento de evaluar la traducción de esta palabra en el campo dos.
- **Campo 2 de palabra:**
Se debe traducir la palabra con ayuda del programa. En este campo se puede ver si existe variación en la respuesta con respecto al campo uno. Esto se usa para analizar si el software tiene impacto en la respuesta.
- **Campo 3 de palabra:**
Comentarios generales de la palabra, relacionados al programa. Se espera recibir retroalimentación del funcionamiento y de la utilidad de la respuesta entregada para la palabra en particular. En el caso de que haya usuarios que tengan un nivel intermedio o

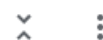
mayor del mapuzugun, también se pueden ver comentarios o dudas relacionadas a la gramática.

Se debe aclarar que existen 6 variaciones de esta parte, debido a que se tienen 6 conjuntos de palabras para traducir, para ver una mayor diversidad de palabras y analizar su resultado, en la Figura 3.11. se puede ver un ejemplo de pregunta.

Parte 2:

En esta parte el usuario debe traducir un extracto de un texto en Mapuzugun, decidiendo libremente si es que usa o no la herramienta para apoyarse. Además, se consulta el grafemario que más le acomoda, de los tres principales (Azümchefe, Ragileo, Unificado), para mostrar el texto a traducir.

Taï Ayewe Foro (Azümchefe)



Aquí te dejaremos el texto para Traducir: (INTENTAR HACER LO MÁXIMO POSIBLE DE PALABRAS), este texto es la primera parte (no concluye)

Pichikalu iñche, amukefun chillkatuwe ruka mew. Fill antü kom chi pu pichikeche chillkatufegelu zewmakefuiñ kakewme ruka küzaw taiñ zoy kimaël fentxen zugu: wirin, zugulwirilün, rakin ka pürun kütü. Kiñe antü mew, chi kimelfe feypieñmew taiñ wirintukuafiel taiñ zoy ayikenoe! taiñ kalül mew.
Llitufun taiñ rakizuamael, ka perumefun taiñ zoy ayikenoe! taiñ kalül mew. Fey, taiñ ayewe foro gefuy, rakizuamkefun taiñ ayewe foro ñi rume fütxa gen mew. Tukulpafun kiñeke mew taiñ ayentunmaetew taiñ ayewe foro. Kiñeke pichikeche zoy alhükelu iñche mew, mara foro nieymi pikefenew egün, feymew yewekefun. Feychi antü amuletulu taiñ ruka mew, rakizuamfun chungelu alhüken taiñ foro, kümeafuy zoy pichi foro niefuli taiñ ayentunoaetew, piwün. (...)

Figura 7.2. Ejemplo del texto a traducir en Azümchefe.

- **Campo 1 de traducción:**

En este campo se debe poner la traducción del extracto mostrado en español. Este campo se evaluará en conjunto con los demás, pero en caso de ser necesario, se verá si su traducción es correcta.

- **Campo 2 de traducción:**

En este campo se consulta si es que se utilizó la herramienta. Con el fin de saber si la traducción resultará útil en la evaluación del software.

- **Campo 3 de traducción:**

Esta es una pregunta relacionada a para que cantidad aproximada de palabras se utilizó la herramienta. Las respuestas posibles son:

- a. Casi nada
- b. Poco
- c. Aproximadamente la mitad
- d. Más de la mitad
- e. Casi todo

En este campo también se obtiene información de cuán útil fue la herramienta, dependiendo de cómo está la traducción.

- **Campo 4 de traducción:**

El campo final de la parte de traducción es para recibir comentarios generales sobre la utilidad de la herramienta en el contexto de traducción de un texto, en la Figura 3.12 se puede ver un ejemplo de esto.

Parte 3:

Está relacionada al funcionamiento del software y la opinión que tienen los usuarios sobre su utilidad y facilidad uso.

- **Campo 1 evaluación final:**

La primera pregunta es si es que se lograron ver las separaciones de las palabras analizadas. Un tema de funcionamiento. Además, en el sentido de la usabilidad se evalúa que efectivamente la persona logre introducir una palabra.

- **Campo 2 evaluación final:**

Este campo pregunta si es que se logró ver el análisis de la palabra segmentada. También en esta se considera un tema de funcionamiento y de usabilidad para llegar a la sección de análisis de las palabras.

- **Campo 3 evaluación final:**

En esta parte, se pregunta por errores del programa. Si es que dejó de funcionar o si es que no entregó lo que debía. En esta sección no se consideran los errores de diccionario, es decir, no es un error para esta parte si es que no encontró una palabra.

- **Campo 4 evaluación final:**

Pregunta sobre “qué te gustó más sobre la herramienta”. Para ver qué aspectos se pueden potenciar en las próximas versiones.

- **Campo 5 evaluación final:**

Pregunta sobre “qué te gustó menos sobre la herramienta”. Para ver qué aspectos se pueden mejorar o directamente eliminar en las próximas versiones del software.

- **Campo 6 evaluación final:**
Evaluación con nota de uno a cinco sobre la dificultad de usar el programa. De 1 a 5, donde 1 es lo más Fácil y 5 lo más difícil.
- **Campo 7 evaluación final:**
Evaluación con nota de uno a cinco sobre la dificultad de hacer las traducciones de palabras . De 1 a 5, donde 1 es lo más Fácil y 5 lo más difícil.
- **Campo 8 evaluación final:**
Evaluación con nota de uno a cinco sobre la dificultad de hacer la traducción del extracto de texto. De 1 a 5, donde 1 es lo más Fácil y 5 lo más difícil.
- **Campo 9 evaluación final:**
Evaluación con nota de uno a cinco sobre la parte visual de la herramienta. Con 1 la peor nota y 5 la mejor nota.
- **Campo 10 evaluación final:**
Evaluación con nota de uno a cinco sobre el funcionamiento general de la herramienta. Con 1 la peor nota y 5 la mejor nota.
- **Campo 11 evaluación final:**
Se pregunta sobre qué tan probablemente se recomendaría esta herramienta a otra persona que está enseñando o aprendiendo mapuzugun. Las opciones son:
 - a. No la recomendaría
 - b. Poco probable
 - c. Muy probable
 - d. Depende de la persona
- **Campo 12 evaluación final:**
En esta última pregunta se esperan comentarios generales de la herramienta por completo.

Debido a las complicaciones para obtener respuestas en la primera encuesta, por el largo y dificultad de esta, se realizó una segunda encuesta, la cual es más corta y agrega una pregunta a los encuestados, respecto de su nivel de mapuzugun.

7.1.2 Segunda Prueba.

La principal diferencia de esta prueba con la anterior es su extensión, la cantidad de palabras en la parte inicial a traducir se redujo desde 9 a 3 palabras para que no sea tan tedioso para los usuarios, tomando en cuenta que usuarios comentaron eso dentro de la primera encuesta.

Además, en la parte 2 se quitó el texto y se reemplazó por una frase debido, principalmente, a que se consideró que el tiempo que se debía invertir en responder la primera encuesta, era excesivo. En el caso de la tercera parte se mantuvo completamente excepto en lo referido al texto, que cambió a referirse a frase. Finalmente, se hizo una sola prueba, a diferencia de la anterior, que tenía 6 variaciones de las palabras de la parte 1.

7.2 Pruebas del apoyo lingüístico

7.2.1 Primera Prueba.

Para el caso de las traducciones de palabras para las personas que utilizaron la herramienta para traducir (descontando expertos y el intermedio que suman 4), se puede ver que existen mejoras en las traducciones, en este sentido de los 4 que respondieron preguntas con la herramienta, 3 no respondieron ninguna pregunta sin ayuda, tuvieron 7, 1 y 4 palabras traducidas con el programa, las cuales tienen como promedio (con mayor puntuación 1 y menor 0), un 0.56, 0.5 y 0.875. Esto significa que no sabían ningún término de los mostrados, pero que con ayuda del programa pudieron ver al menos el sentido de las palabras que estaban traduciendo. Además, aparece un participante que hizo 4 palabras sin ayuda y sumó una más con ayuda puntuando la nota máxima (1).

Con respecto a la evaluación del texto a traducir, de las 9 personas, 7 utilizaron la herramienta, pero de estos, 4 la utilizaron “muy poco”, 1 la utilizó “más de la mitad” y 2 la utilizaron “casi toda” en la traducción.

De las 4 personas que la utilizaron muy poco, no se puede sacar información sobre la herramienta debido a que no fue determinante al momento de traducir. Pero en el caso de las personas que la utilizaron para casi todas las palabras, hicieron aproximadamente la mitad del texto y palabras correctas, esto versus su rendimiento sin traductor, que está rondando 0. En este sentido, existe una ayuda a una persona que tiene poco conocimiento del Mapuzugun.

Sin embargo estos resultados, si bien entregan información acerca de las traducciones, no entregan un resultado que pueda ser analizado en profundidad, tomando en cuenta que son pocas respuestas y, muchas veces, las palabras a traducir varían de persona a persona.

7.2.2 Segunda Prueba.

En esta prueba, al ser menos palabras, se logró que la mayoría de las personas contestara la encuesta, a diferencia de la primera. Además, se pueden sacar ciertos datos que permiten afinar el público al que le resulta de mayor utilidad la herramienta. Las notas de esta sección son de 0 a 3

en donde 0 es una traducción completamente errónea, y 3 es una buena traducción sin errores (aunque no son respuestas únicas).

Grupo 0: Este grupo es el correspondiente al que tiene menos conocimientos de la lengua, se puede ver que en la primera palabra (“elukelafimu”) se mantienen las puntuaciones en 2, no existe ningún cambio de hacerlo con o sin la herramienta. Las siguientes dos palabras tienen un aumento de puntuación al usar la herramienta, para la palabra “pemurpayafuyu” se pasa de un 1 a un 3, mientras que “kujinerkeeiñmu” pasa de 0 a 1.33. La frase, al igual que la segunda palabra, pasa de tener promedio 1 a tener promedio 3. En 3 de 4 preguntas, se ve un aumento de rendimiento de los usuarios al usar la herramienta.

Grupo 1: Este grupo consta de personas que pueden decir ciertas frases en Mapuzugun, pero que aún no pueden entender una conversación. La puntuación, en el caso de “elukelafimu”, baja y pasa de un 2 a un 1.6. Luego, “pemurpayafuyu” tiene un aumento de 1 a 2, “kujinerkeeiñmu” de 0.67 a 1.4 y la frase de 2 a 2.25. En este caso, existe una baja del rendimiento en una de las 4 preguntas y en las demás un aumento, aunque no tan grande como el del grupo 0.

Grupo 2: Se podría definir este grupo como un intermedio, debido a que puede entender las conversaciones en Mapuzugun, aunque no pueda participar activamente de estas. En “elukelafimu” al igual que el grupo anterior, se ve un retroceso, pero tanto en “pemurpayafuyu” como en “kujinerkeeiñmu”, se ve un aumento de 1.83 a 2.12 y de 1.75 a 2.0, respectivamente. La frase por su lado también aumenta su puntuación.

Grupo 3: Este es el grupo que corresponde a estudiantes avanzados y personas que podrían ser consideradas Neo-hablantes o estar cerca de esa categoría. En ese sentido, se tienen sin ayuda varios puntajes altos: 2.75 en “elukelafimu”, 2.33 en “pemurpayafuyu” u 2.43 en “kujinerkeeiñmu”. Pero, además, a diferencia de las 2 anteriores, si existe un crecimiento en la primera palabra, sumado a que también se mejora la puntuación en la segunda y tercera. Además, en el único caso que no aumenta es en el de la Frase, porque cada una de las personas de este grupo lo contestaron de forma correcta y por lo tanto no se puede advertir el efecto de la herramienta.

Grupo 4: Para el grupo de los hablantes se tiene una particularidad, estas 3 personas, no utilizaron la herramienta según sus respuestas. Por lo tanto solo se tiene puntuación de las palabras originales. Esta puntuación es “elukelafimu”, 2.0, “pemurpayafuyu”, “kujinerkeeiñmu”. la frase también tiene una buena puntuación 2.67.

Como resultado general, es decir, juntando todos los grupos, “elukelafimu” tiene una baja de 2.33 a 2.14, al usar la herramienta, “pemurpayafuyu” tiene un aumento de 1.94 a 2.44, “kujinerkeeiñmu” tiene un aumento de 1.81 a 2.05 y la frase de 2.42 a 2.72.

Estas mejoras se pueden deber a que se aporta información morfológica que puede hacer sentido al usuario, además el grupo 2 y 3, que tienen conocimientos, pero que se encuentran en nivel intermedio o avanzado, es donde se muestra más avance de hecho el único grupo donde “elukelafimu” no tiene un retroceso es el grupo 3.

Esto nos puede mostrar que al tener una base relativamente sólida de Mapuzugun, este software aporta datos que son relevantes para una traducción un tanto más precisa. No así en el caso de los hablantes, que no la necesitan para obtener buenos resultados o que se sienten seguros de sus traducciones. Se destaca que, si bien existe un grupo que afirma “Todavía no se mucho mapuzugun”, de igual manera se pudo tener para estas 6 personas un puntaje 3 (máximo) tanto para “pemurpayafuyu” como para la Frase y por lo tanto, aún casi sin base esta herramienta podría entregar información para poder esbozar la definición de la palabra.

Además se puede decir que “elukelafimu” puede tener problemas debido a que es una palabra que es más fácil de acceder en niveles iniciales de Mapuzugun y por lo tanto más conocida por los usuarios, debido a esto el programa pudo haber confundido a ciertas personas en este punto.

El resumen de estos resultados está en la siguiente Tabla 7.1:

Palabra	Grupo 0	Grupo 1	Grupo 2	Grupo 3	Grupo 4	General
Elukelafimu s/a	2	2	2	2.75	2	2.33
Elukelafimu c/a	2	1.6	1.86	2.86	-	2.14
Pemurpayafuyu s/a	1	1	1.83	2.33	2	1.94
Pemurpayafuyu c/a	3	2	2.17	2.83	-	2.44
Kujinerkeeiñmu s/a	0	0.67	1.75	2.43	3	1.81
Kujinerkeeiñmu c/a	1.33	1.4	2	2.86	-	2.05
Frase s/a	1	2	2.57	3	2.67	2.42
Frase c/a	3	2.25	2.71	3	-	2.72

Tabla 7.1. Resumen resultados parte lingüística segunda prueba.

7.3 Pruebas de usabilidad.

Para las pruebas de usabilidad del software, se tuvieron las mismas dificultades para encontrar los grupos de usuarios que se explicaron antes.

7.3.1 Primera prueba.

Dentro del ámbito de la usabilidad, en primera instancia se preguntó si es que se logró ver la separación de las palabras. De las personas que respondieron la encuesta, todas dijeron sí las habían visto.

Luego, sobre la pregunta sobre el poder ver la explicación de la separación, solo una persona afirmó que no pudo hacerlo, pero en el final de la encuesta comenta que llegando casi al final de esta, logró ver que se podía desplegar la explicación.

Las siguientes preguntas la evaluación en notas de cinco aspectos del software, los resultados de esta parte son los promedios de las 9 personas que contestaron la encuesta, en cada pregunta se podía escoger entre 1 y 5, depende de la pregunta a qué respuesta corresponden esos números:

- Dificultad de utilizar el programa, siendo 1 fácil y 5 difícil, tuvo un promedio de 2 .
- Dificultad de traducir una palabra, siendo 1 fácil y 5 difícil, tuvo un promedio de 2.6 .
- Dificultad de traducir un texto, siendo 1 fácil y 5 difícil, tuvo un promedio de 2.8 .
- Sobre la parte visual se tiene un promedio de 2.8, siendo 5 la nota más alta, y
- Sobre la nota general de la herramienta, quedó en un 3.8 .

Las últimas dos preguntas corresponden a la posible recomendación de la herramienta, todas las personas manifestaron que es muy probable (seis personas) o que depende de la persona (3 personas) que se recomiende el uso de este programa a otras personas que estén aprendiendo o enseñando mapuzugun.

La parte de comentarios, tanto de las partes de traducciones y/o de la parte de usabilidad se verán en la siguiente sección en la que se hará el análisis para la mejora del software.

7.3.2 Segunda prueba.

Las preguntas para la segunda versión de la encuesta no cambian para esta sección, por lo que en el ámbito de la usabilidad, se tienen dos encuestas iguales, por lo que para analizarlo se debe tener en cuenta, primero que la interfaz cambió, y que, además, se actualizó el analizador morfológico. Sobre esto último se puede revisar la sección 6.4.

La primera pregunta es si es que se logró ver la separación de las palabras. De las 32 personas consultadas, 6 afirman que no se pudo ver.

Sobre si vieron la explicación de la separación, 7 afirman que no pudieron ver la explicación, 6 son las personas que no lograron ver la separación de palabras. Las otras 25 personas, lograron ver tanto la separación de las palabras como la explicación de las partículas que se utilizan en ella.

En la sección de evaluación numérica para cinco aspectos del software, se pueden ver los resultados para las 32 personas que contestaron la encuesta, en cada pregunta se podía escoger entre 1 y 5, depende de la pregunta a qué respuesta corresponden esos números:

- Dificultad de utilizar el programa, siendo 1 fácil y 5 difícil.
- Dificultad de traducir una palabra, siendo 1 fácil y 5 difícil.
- Dificultad de traducir un texto, siendo 1 fácil y 5 difícil.
- Sobre la parte visual 5 corresponde a la nota más alta, y
- Sobre la nota general de la herramienta, también 5 corresponde a la nota más alta.

	Dificultad uso	Dificultad traducción de palabra	Dificultad traducción frase*	Nota parte Visual	Nota general
Prueba 1 (referencia)	2	2.6	2.8 (texto*)	2.8	3.8
General	2.33	2.87	3.07	3.63	4.0
Grupo 0	2.67	3.17	3.33	3.83	3.17
Grupo 1	2.71	2.86	3.71	3.29	4.0
Grupo 2	1.86	3.0	2.71	4.14	4.71
Grupo 3	1.86	2.43	3.14	3.86	4.29
Grupo 4	3	3	1.67	2.33	3.33

Tabla 7.2. Resultados de la evaluación de usabilidad.

En general, todos los resultados tuvieron un aumento, tanto la dificultad de uso del software, como la evaluación en general y los aspectos visuales de la herramienta.

Estos resultados se muestran en la Tabla 7.2, agrupados por grupo. La referencia a general, implica valores obtenidos en conjunto de todos los grupos.

Analizando la encuesta, se puede ver que los grupos 2 y 3, que tienen un mapuzugun de intermedio hacia arriba, son los que menor dificultad encuentran para usar el programa, contrastando con un 3 de dificultad del grupo de hablantes, mientras que, en un punto intermedio, se pueden ver los grupos 0 y 1.

Para la dificultad de traducir palabras, se mantiene el grupo 3 como el valor más bajo y los hablantes siguen teniendo valores más altos. En la dificultad para traducir frases, los hablantes bajan su número junto con el grupo 2 y todos los demás aumentan.

En general en promedio, todas las dificultades subieron, pero existen grupos específicos como hablantes, el grupo 0 y el grupo 1, en particular que tienen un puntaje mayor y, en particular, el grupo 2 y 3 que bajo esos números en algunos casos.

Con respecto a la parte visual, se tiene un un aumento significativo de 0.8 puntos, esto se condice con el cambio de interfaz, influyendo quizás en el aumento de 0.2 puntos en la evaluación general.

La pregunta sobre si se recomendaría la herramienta, muestra que 17 personas la recomendarían muy probablemente, 8 lo harían dependiendo de la persona a la cual se le recomendará. Además, la opción poco probable salió 3 veces y la que no se recomendaría salió 2 veces. Esto no suma 32 debido a que ninguna de las preguntas de esta encuesta es obligatoria y, por lo tanto, hubo dos personas que no respondieron esta pregunta.

La parte de comentarios, tanto de las partes de traducciones y/o de la parte de usabilidad se verán en la siguiente sección en la que se hará el análisis para la mejora del software.

7.4 Los comentarios.

7.4.1 Primera prueba.

En esta parte, se seleccionaron solo los comentarios que tienen que ver con el programa, tanto en sus fortalezas como en sus debilidades, los comentarios que tienen que ver con la opinión de la existencia del software se dejaron fuera, al igual que comentarios del tipo felicitación.

Como se mencionó se pueden dividir los comentarios, en partes técnicas, lingüísticas y además en fortalezas y debilidades. Estas categorías ayudan a hacer un panorama completo de lo que se aportó por parte de los usuarios.

Dentro de lo visual se pueden ver comentarios sobre la interfaz, que es confusa y que se requiere un trabajo mayor sobre ella debido a que tiene un aspecto muy simple o “sin esfuerzo”. Esta interfaz como se puede ver en la Figura 6.1, un usuario propone colores para separar los morfemas, para que se entienda su correspondencia con la explicación respectiva.

Aunque, surgieron comentarios como “el programa es amigable”.

Algunos usuarios notan que existen ciertas palabras que no funcionan al probarlas y también se debe ver el porque pasa eso y trabajar sobre los posibles errores. Otro comentario de las palabras, fue que si bien se pueden entender los morfemas, hay casos en que no es suficiente para hacer la interpretación del significado en castellano de la palabra en cuestión. Sin embargo, se toma como un buen ejercicio pedagógico el tener la separación de las palabras en general. Otros comentarios positivos van en el sentido de la mayor rapidez que se tiene en el programa, comparado a tablas o diccionarios que suelen ser consultados para el estudio del Mapuzugun.

También existen comentarios del tipo que se tienen muchas opciones para ciertas palabras y por lo tanto se requiere hacer una poda para las separaciones que podrían ser redundantes, esto confundió a los usuarios al punto de no poder escoger una alternativa o no poder interpretar la traducción. Lo anterior se puede explicar, en parte, en que existen ciertos morfemas que se pueden utilizar de distinta forma, dependiendo del contexto y las palabras que los acompañan. Un ejemplo sería el caso de “-en” que podría utilizarse en contextos imperativos y/o indicativos.

También la redundancia se da en que existen ciertas unidades que son fusiones de dos morfemas, lo que como se explicó, se da porque se tiene, para este programa, definiciones comunicativas más que lingüísticas y, por lo tanto, esta fusión de morfemas podría cambiar el sentido de los morfemas por separado. Aplicado a este último comentario, se hizo notar la ausencia de la combinación de los sufijos -a- y -fu- que en conjunto (-afu-) generan una frase hipotética y por separado modifican el tiempo de esta.

Separador de Palabras Hola! Aquí te dejo una herramienta para que te ayude a aprender más Mapuzugun, espero que con esto entiendas más este idioma (el mejor).	Nhemül Wüzalkawe Mari mari! Ta mu elkünul-ae-yu kiñe küzawpeyüm tami kelluaetew zoy kim mapuzugual, pewmagen mai tüfa mu zoy azümneaimi tüfachi zugun llemai (tati zoy kümelu).
---	--

Palabra / Nhemül | Aquí pon tu palabra /tüfa mu toku | Separalo / Wüzalkafige

Así lo separé | Femgechi wüzalkafin

amulen

Creo que está escrito en Azümche | Wirintukugelu Azümche mu txokifin

► -amul-en
▼ -amul-en

amul	en
vtr enviar, echar a andar, encaminar; hacer	de tú a yo "orden"

► Expresion regular (Para computación)

► -amu-l-en
▼ -amu-l-en

amu	l	en
vi irse, andar, continuar, avanzar	hacer hacer algo, hacerle algo	de tú a yo "orden"

► Expresion regular (Para computación)

► -amu-le-n

Comentario / Tami chumgentufiel tati küzawpeyüm

Aquí escribe como podría ser mejor nuestra herramienta |
 | tüfa mu wirintukufe chumgechi zoy kümelefel tati küzawpeyüm

Envíalo / Werkülfige

Ejemplos/Azümam

Aquí hay algunas palabras de ejemplo para que entiendas como usar esta herramienta

Tüfa mülei kiñeke nhemül tami azümafiel chumgechi püneafiel tüfachi küzawpeyüm

Wazküy, que significa "hirvió" se podría separar como:

► -wazkü-y
- **Amuaiñ**, que significa "iremos" se podría separar como:

► -amu-a-ñ
ayüwkülen, que significa "estoy contento/feliz" se podría separar como:

► -ayüwküle-n
► -ayüw-küle-n

También, Aquí se dejará un texto para que saques palabras y las uses en la herramienta. **La traduccion a veces no es 100% literal**	Ka femgechi, Tüfa mu elkünugeal kiñe wirintukun tami entual nhemül tami püneafiel küzawpeyüm mu
Hola hermanos hermanas. Yo me llamo Juan. Conversaremos de mi descanso de invierno (vacaciones). Yo fui al sur, por Gorbea. Fui a estudiar en un koneltun*(internado de mapuzugun). Yo voy a temuco habitualmente, pero este año no fui, en el verano iré a ver a mi familia a temico. Cuando estaba en Gorbea, hable con algunas personas, ell@s también estudian mapuzugun. Otras personas saben hablar mapuzugun. Ahora yo se más mapuzugun, así que todavía estoy feliz	Mari mari pu peñi pu lamgen. Ñche Juan pigen. Nütxamkawaiñ tañi pukem ürkütun mew. Ñche amun willi mew, gorbea püle. Chillkatumen kiñe koneltun mew. Ñche amuken Temuco mew, welu tüfachi txipantu amulan, walüg mew pemean tañi reñima Temuco mew. Mülelu gorbea mew, Nütxamkan kiñeke che inchiñ, feyegün ka chillkatukeigün mapuzugun. Kakelu che kim mapuzugukelu. Fanten mew ñche zoy kimün mapuzugun, feymew petu ayiüwkülen.

Figura 7.3. Captura de pantalla de la primera versión del software.

Al responder la pregunta “Qué es lo que más te gustó del programa”, se repite el aporte que hace la separación por morfemas, además se hace mención a la detección de grafemarios.

Como paréntesis, se plantea en un comentario en particular que existe un nivel mínimo de Mapuzugun desde el cual se puede aprovechar mejor esta herramienta, pero no se indica de forma concreta.

Al preguntar por lo que menos gustó, destaca la parte visual, tanto por lo poco llamativa, como por la disposición de la información. Otros comentarios tienen que ver con la repetición y con las definiciones de los morfemas. Además, se hace mención a que se podría entregar directamente una traducción al castellano.

También fue propuesto por un usuario la creación de un glosario de las definiciones de las unidades morfológicas que se presentan.

7.4.2 Segunda prueba.

Los primeros tres comentarios tuvieron problemas técnicos, porque la segunda palabra tenía un error de tipeo en la encuesta, se puso “perpamuafuyu” en vez de “pemurpayafuyu”. Luego de eso se tiene una prueba más estable en el sentido de la construcción de la encuesta. Esta encuesta también se mostrará dividida en grupos dependiendo de la autoidentificación de los participantes de acuerdo a su nivel de Mapuzugun.

Grupo 0: En este grupo, los comentarios de las traducciones tienen que ver con que no se pueden lograr las traducciones, incluso entendiendo cada parte por sí sola de la separación.

En lo que más les gustó a este grupo se encuentra la característica de poder ver la composición de las palabras en Mapuzugun gráficamente. Complementando con lo que menos gustó, existen ciertas formas de describir palabras, como VTR, S, VI, que indican el tipo de palabra, que no son de conocimiento general, incluso sabiendo que es verbo o sustantivo, las siglas no necesariamente son conocidas.

Luego, en los comentarios generales se puede ver que se agradece la herramienta, pero que quizás la dificultad para traducir a partir de la herramienta puede ser muy alta, asociado, según un usuario, a que podría faltar una base en el idioma. Además, para una persona no fue útil, por no entender su funcionamiento.

Grupo 1: Uno de los participantes de este grupo dice que está bien especificado el significado de las partículas. Además, surgen varios comentarios acerca de que personas principiantes en el idioma podrían tener problemas para su uso. Esto se suma a que existen usuarios que no entendieron que había un botón para cambiar de idioma Mapuzugun a Español dentro de la

herramienta. Y relacionado a esto último, se da a entender que, sobre todo si no se hace el cambio de idioma, la herramienta no sería lo suficientemente intuitiva para este nivel.

Lo que más gustó en este grupo también fue el hecho de que se muestre la descomposición, y que hayan varias opciones. Una persona también notó como algo positivo que esta herramienta esté en Mapuzugun.

Luego, en lo que menos gustó, se comenta que deberían haber ejemplos, a pesar de que existen, quizás no están en un lugar fácil de encontrar. La falta de instrucciones y la dificultad de entender el funcionamiento para “no hablantes”. Y relacionado a esto último, se pide que los botones Mapuzugun y Español estén más visibles. Otro comentario es que no entrega el significado final de la palabra. Finalmente, un usuario piensa que quizás la explicación puede estar muy sujeta a la interpretación de las personas.

En comentarios generales, se hace nuevamente la acotación de un posible problema con las personas que no conocen las partículas o morfemas y se sugiere un glosario, además de la incorporación de audios.

Grupo 2: Este grupo fue más activo en comentarios, en el caso de las traducciones, se hace referencia a un morfema duplicado, estos morfemas son iguales pero que tienen distintas funciones comunicativas, esto podría ser mejorado fusionándose ambas partículas, pero se debe hacer notar que podría estarse refiriendo a dos cosas distintas. También, se comenta que a veces no basta el hecho de tener las partículas para poder hacer la traducción y que de preferencia de los usuarios sería bueno mostrar ambas cosas (separación y traducción). Además, se hace notar que algunas palabras que no son verbos son separadas y que, quizás, eso sería un error en ciertos casos.

Dentro de lo que más le gustó, está el hecho de la separación y su explicación, además comentan que se puede practicar conjugación.

Lo que no gustó fue que quizás podría tener “ más palabras en español” debido a que hay gente que podría no entender que este programa tiene una versión en castellano, por no ser intuitiva la forma de acceder a ese contenido. Además pensar en las palabras que sin contexto se pueden ver de varias formas y esas variaciones podrían confundir a ciertas personas sin práctica o conocimientos. También se nota la falta que hace una traducción directa a este trabajo debido a que los usuarios también ven como algo que no les gusta de este trabajo ese punto.

Finalmente en los comentarios generales, se hace énfasis en que algunos colores no tienen el contraste suficiente cuando se muestran las explicaciones de los morfemas. Un usuario también sugiere una aplicación móvil.

Grupo 3: Los comentarios en la parte de las traducciones tienen que ver con posibles errores gramaticales. Además de que existen algunos que también hablan de que los estudiantes iniciales se podrían confundir con la cantidad de opciones que aparecen para ciertas palabras. Estos

participantes dieron comentarios en general sobre los morfemas, no solo sobre el funcionamiento, apareciendo algunos quienes dijeron que el programa les había ayudado a entender alguna partícula.

También que los usuarios deben tener un nivel mínimo para usar la herramienta, o que muchas veces hace falta un contexto.

Lo que más le gustó a este grupo, tiene que ver con la utilidad en palabras extensas o de mucha aglutinación, aportando en el aprendizaje de la gramática. Se hizo un buen comentario sobre la interfaz y sus colores. Y un comentario dice que le gusta que requiera un esfuerzo y no entregue la traducción directamente.

Pasando a lo que menos gustó en este grupo, algunas personas nombraron que se podría incluir la traducción, además los morfemas no siempre son claros para las personas que los leen y por lo tanto se podría tener algún glosario para que se pudiera profundizar. También que algunas traducciones podrían ser muy abiertas a la interpretación y poder confundir su significado dentro de la palabra en cuestión, también se habla de ciertas combinaciones de partículas que tienen un significado en conjunto que no están incluidas de esa forma en esta versión. También sobre gramática, se pueden tratar de incluir versiones “incorrectas” de palabras de entrada.

Finalmente en comentarios generales, se celebra la idea pero no se hacen comentarios relevantes a la mejora o cambios en el programa.

Grupo 4: En este grupo no se hicieron mayores comentarios, uno de estos que hay typos en la traducción de un morfema y que, además visualmente, se hace la acotación de que algunos colores no se distinguen y que en ciertos casos en las separaciones de la palabras muy largas se queda corto el espacio para mostrar y se pueden perder algunas palabras o confundir a las personas.

7.5 Modificaciones hechas a partir de la retroalimentación.

A partir de los comentarios se modifican algunas definiciones que contienen errores, se agrega la fusión de morfemas AFU y PELA, se eliminan ciertas palabras que generan conflictos o redundancias como en el caso de GE o de palabras duplicadas (sufijo TU por ejemplo). Además, se hace un trabajo de mejora de la interfaz de la página, para que esta sea más atractiva, más dinámica y pueda tener apoyo visual para las partes más didácticas. También se separan los contenidos en pestañas, para que no esté todo en una sola página. La interfaz se puede ver en las Figuras 6.2, 6.3, 6.4 y 6.5.

Un cambio posterior a la interfaz, fue el hecho de agrandar los botones de idiomas e invertir su orden para que el de Español se vea primero y luego el Mapuzugun.

Además se incluyeron algunas reglas sobre los modos del mapuzugun que no se pueden mezclar dos modos, relevante principalmente para el caso de las negaciones, y se cambiaron las formas de algunos morfemas, como el “-ymi”, cuya expresión regular era $r'[iy]\{0,1\}mi'$ por simplemente $r'[iy]mi'$ esto debido a que, si bien existe un caso cuando se juntan los morfemas “fi” e “ymi” resultando “fimi” (desapareciendo la Y en este caso), esto traía confusiones con el modo imperativo el cual utiliza el morfema “mi” generando duplicados para ciertas palabras.

The screenshot displays the updated Mapuzugun interface. At the top, a blue header bar contains the title "Nhemül Wüzalkawe" and two buttons: "Mapuzugun" (highlighted in green) and "Español". Below the header, a navigation bar shows three tabs: "Küzawpeyüm" (active), "Chumgechi püneafiel", and "Gütxam püneafiel tufa mu". The main content area features a large grey box with a white cross icon and the text "Mari mari!" followed by a paragraph in Mapuzugun. Below this, there is a section labeled "Nhemül" with a text input field containing "Tufa mu tukulfe tami nhemül" and a green "Wüzalkafige" button. Further down, another section labeled "Tami chumgentufiel tati küzawpeyüm" contains a text input field with "Tufa mu wirintukufe chumgechi zoy kümeleafel taiñ küzawpeyüm" and a green "Werkülfige" button.

Figura 7.4. Página principal de la interfaz actualizada.

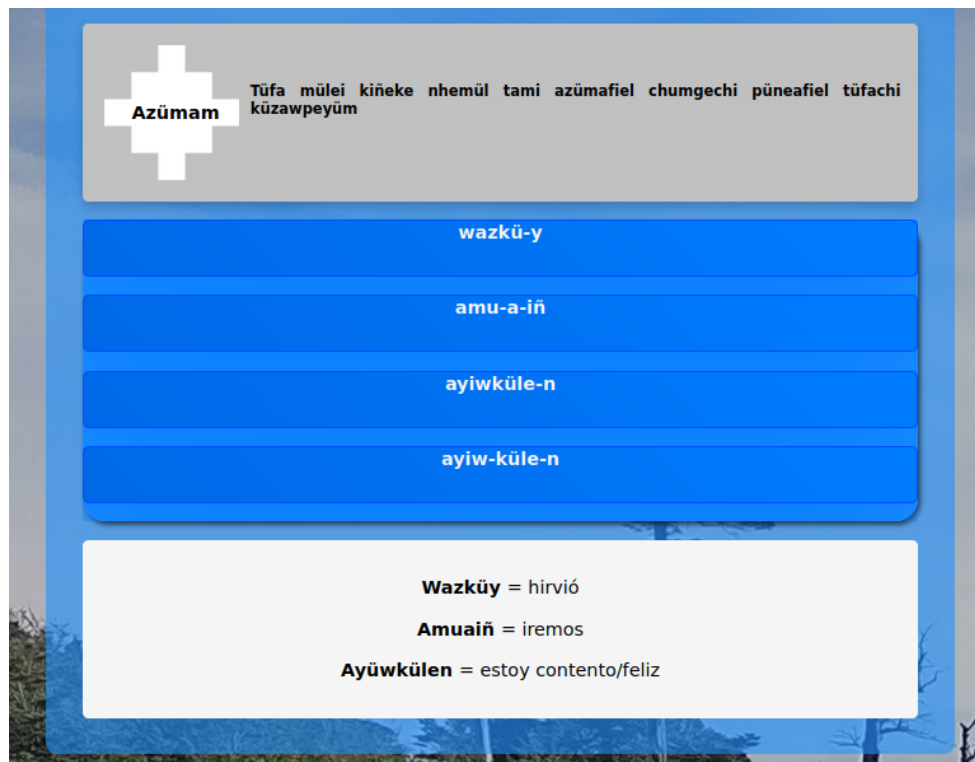


Figura 7.5. Página de ejemplos de uso de la herramienta de la interfaz actualizada.



Figura 7.6. Texto para probar la herramienta, en la interfaz mejorada.

Nhemül

Tüfa mu tukulfe tami nhemül

Wüzalkafige

AMUAFUIÑ

Wirintukugelu Azümchefe mu txokifiñ

amu-afu-iñ

amu	afu	iñ
vi lrese, andar, continuar, avanzar	hipotético	Nosotros (mas de 2 personas)

► Expresion regular (computación)

Tami chumgentufiel tati kūzawpeyüm

Tüfa mu wirintukufe chumgechi zoy kúmeleafel taiñ kūzawpeyüm

Werkülfige

Figura 7.7. Nueva interfaz con resultados.

8 Conclusiones y Trabajo Futuro.

8.1 Conclusiones.

En esta tesis, se han implementado un traductor de grafemarios, un detector de grafemarios, un analizador morfológico y un traductor informal, los cuales se han integrado y puesto a disposición para su utilización por personas que están relacionadas con el aprendizaje de idioma Mapuzugun. Estas herramientas son presentadas a los usuarios a través de una interfaz lo que facilita su uso. Además, se hicieron pruebas con usuarios para validar la herramienta y proponer modificaciones a futuro.

Uno de los cambios más importantes con respecto a lo que se tenía planeado en un inicio, es que se pensaba adaptar un analizador morfológico existente, pero finalmente, por razones de facilidad para los usuarios que no son expertos en lingüística, se optó por generar uno que estuviera más enfocado en la comunicación que en la gramática. Debido a esto se empezó desde cero y se pudo llegar a un resultado que separa las palabras de una forma más parecida a la enseñanza del Mapuzugun y, alejándose de un enfoque de análisis más técnico/teórico de la lengua.

En el análisis de los resultados presentados en la sección anterior, se puede ver que existe un grupo de personas a las cuales les fue de más utilidad: aquellas que se auto-definieron como "He estudiado o he ido a cursos. Y puedo entender conversaciones de otras personas en Mapuzugun." y "He estudiado o he ido a cursos. Puedo mantener conversaciones en Mapuzugun.". En estos grupos surgieron las mejores evaluaciones, así como también, la mayor cantidad de comentarios respecto a que se puede mejorar. Esto está relacionado con el hecho que las dudas y preguntas sobre la gramática surgen a partir de un conocimiento inicial de Mapuzugun, pero donde ya se tiene una base y se necesita pulir los detalles de formas, tiempos y énfasis que pueden entregar los sufijos a un verbo.

Si bien existen ciertos comentarios de los usuarios respecto a la dificultad de uso, la evaluación del software muestra que en las palabras difíciles y largas (como la 2 y 3 de la prueba), se logra una mejor comprensión de estas a nivel general. Esto se puede verificar en la puntuación que se obtiene con y sin el programa en esas palabras para la encuesta, por otro lado existen ciertas palabras menos avanzadas, con respecto a cómo se enseña el Mapuzugun, que al estar más consolidadas en las personas, podrían generar confusiones ver algunas separaciones que los usuarios no conocen.

Un tema del aprendizaje al que se debe poner énfasis, es que los hablantes y las personas que están aprendiendo Mapuzugun como segunda lengua, tienen dinámicas distintas a las personas que tienen el Mapuzugun como primera lengua. Por lo tanto se deben generar herramientas focalizadas para este último grupo, por ejemplo para que puedan hacer su vida en Mapuzugun en todos los espacios posibles. Este tipo de separaciones gramaticales no siempre tiene sentido para una persona que aprendió Mapuzugun de forma natural. Pero sí, este tipo de herramientas podrían aportar en la tarea de enseñanza del Mapuzugun, tratando de acercar los pensamientos acerca del Mapuzugun desde estos dos mundos, el de Mapuzugun como primera lengua con el de Mapuzugun como segunda lengua.

Otro aprendizaje de este trabajo tiene que ver con que si bien de la primera a la segunda prueba hubo un incremento en la parte visual, esto no está directamente relacionado con la facilidad de utilizar la herramienta, sino que se debe a haber hecho más explícito temas como los cambios de idioma o las formas de funcionamiento.

Además, se aprende que es necesario complementar este trabajo con materiales de apoyo al uso de la herramienta, tomando en cuenta que ésta forma parte de un contexto en donde el Mapuzugun, como lengua minorizada, no se encuentra con la disponibilidad de los idiomas dominantes, como el Castellano. Por esto mismo, las herramientas por sí solas deben tratar de subsanar en algún grado las carencias de conocimiento de fuentes de palabras, diccionarios, textos ejemplos y, además, la interacción social en Mapuzugun que es prácticamente nula dentro de la vida pública de la sociedad Chilena.

Desde el aspecto lingüístico, se hace notar la falta de fuentes de datos y corpus formales de Mapuzugun. En los idiomas dominantes, o los que tienen corpus más trabajados, las tareas como las que se hacen en este trabajo no son hechas con reglas, si no que se automatizan apoyándose en inteligencia computacional que requiere corpus bien desarrollados.

Una de las conclusiones más importantes tiene relación con el hecho de que esta herramienta sigue estando abierta a mejoras y posibles desarrollos. En el espíritu de continuidad de este trabajo, se dejan a disposición de todas las personas la página abierta² y el código fuente abierto de la aplicación³ para copias locales y colaboraciones tanto dentro como fuera de la academia.

8.2 Trabajo futuro.

8.2.1 Posibles trabajos directamente relacionados.

Este no es ni el primero ni el último trabajo relacionado al Mapuzugun y en este sentido, se dejan tareas posibles para que sigan siendo desarrolladas.

En primer lugar, que el detector de grafemarios desarrollado sea comparado con un detector entrenado en base a datos, para ver si es que se puede tener una ventaja a partir de esto. De esta misma forma, también se podría hacer esta comparación con un analizador morfológico entrenado bajo las mismas lógicas.

También, se podría implementar una identificación de raíces desconocidas y pedir, en caso de que sea posible, una traducción o anotación a esta unidad morfológica.

Complementar esta herramienta con un sistema de anotación de palabras y/o frases para generar un conjunto de datos para posibles tareas futuras ya sea de analizador morfológico, como de una traducción directamente al castellano, en donde se necesiten datos para entrenar modelos de estos.

² <https://crahumadao.pythonanywhere.com/> Link directo a la herramienta online.

³ <http://github.com/crahumadao/kaxvkaam> Repositorio de la herramienta.

También, se debe hacer una revisión y método para ordenar las posibles separaciones, intentando hacer que vayan de lo más probable a lo menos probable, para que los usuarios no se lleven el trabajo de priorizar estas posibilidades por sí mismos en casos de personas que estén en un nivel inicial.

Además, esta herramienta se podría hacer cargo de las palabras de los usuarios integrando audios para que las personas que usen el programa tengan clara la pronunciación en Mapuzugun.

Un trabajo que debe ser continuo, es ir mejorando la interfaz y llevando este software a plataformas móviles como celulares y tablets, y tomando en cuenta, además, las condiciones existentes en ciertos territorios rurales o con problemas de conectividad.

8.2.2 Trabajo futuro no directamente relacionado.

Sí bien este trabajo trata de un aspecto morfológico, de apoyo al aprendizaje y a las traducciones del Mapuzugun, esta no es la única carencia tecnológica de este idioma. Se necesitan desarrollar diversas aplicaciones que puedan llevar al Mapuzugun a todos los niveles y espacios de la computación.

Posibles trabajos para los cuales se podrían obtener datos, de otras fuentes:

- Traductor automático basado en textos bilingües.
- Reconocedores de voz, trabajo extenso debido a la cantidad de variaciones fonéticas del idioma en los distintos territorios. Para este trabajo la CMU tiene un dataset que podría ser utilizado.
- Generadores de voz en Mapuzugun, los cuales sirvan para que los estudiantes y/o hablantes del Mapuzugun tengan herramientas tecnológicas para poder escuchar una lengua que no está presente de forma masiva en redes sociales o medios de comunicación.

Otro tema importante es la localización de los software de uso cotidiano al Mapuzugun y, la generación y necesaria validación, de neologismos relacionados a temas de la web o computadores, para que el Mapuzugun esté presente en esta.

Pero no puede quedar solo en los productos que se surgen desde la computación sino que además se debe incluir la generación de conocimiento de computación en Mapuzugun. Es decir, empezar a desarrollar conceptos de computación, como algoritmos, programas y todo lo relacionado con esta área en Mapuzugun, con el objetivo de enseñar, en un futuro, computación en este idioma. Se recalca que ninguna de estas tareas se puede hacer sin un mínimo conocimiento del Mapuzugun. También se debe tomar como un aspecto central, que todo ese tipo de desarrollos debe ser validado por personas del pueblo Mapuche, ya sea en pruebas de softwares en donde se incluyan participantes Mapuche, como validaciones de sabios y sabias Mapuche, de los conceptos que se insta a empezar a desarrollar.

8.2.2 Trabajo futuro de otras lenguas.

Este software pretende ser un aporte al Mapuzugun, pero no es el único idioma que ha sido minorizado. En todo el país (y el mundo) existen idiomas que, hoy, no son dominantes y se necesita un esfuerzo desde la computación para aportar en la revitalización de estos.

Este software se puede aplicar a idiomas que tengan similitudes al Mapuzugun, ya sea por su contexto de minorizado o por tipología. Se necesitan personas que conozcan el idioma, la gramática y la computación para poder trabajar este tema. Ninguna de estas tres áreas puede quedar fuera, porque las tres aportan las herramientas necesarias para que una herramienta pueda tener un mínimo de validez frente a las personas que son portadoras de las lenguas o que están trabajando para que éstas no se pierdan.

Se necesita que los estudiantes, particularmente de computación, puedan aprender estos idiomas para poder trabajar con ellos, si no se hace, se corre el riesgo de hacer un trabajo o proyecto que no sea utilizado por las personas a las cuales está dirigido, o que tenga un efecto negativo sobre la realidad de la lengua a la que se pretende aportar. Relacionado a este punto, se adquiere una responsabilidad con un idioma al trabajarlo, sobre todo si está en un proceso de intentar salir de la minorización en la que está hoy.

9. Bibliografía.

- [1] Marimán, P., Caniuqueo, S., Millalén, J., Levil, R. (2006) Escucha Winka .Santiago de Chile, Lom Ediciones.
- [2] [Encuesta CEP: 67% de la población mapuche no habla ni entiende el mapuzugun.](https://www.latercera.com/noticia/encuesta-cep-67-de-la-poblacion-mapuche-no-habla-ni-entiende-el-mapuzugun/)
<https://www.latercera.com/noticia/encuesta-cep-67-de-la-poblacion-mapuche-no-habla-ni-entiende-el-mapuzugun/>
- [3] Unesco. (2010). Atlas de las lenguas en peligro del mundo. 3ra edición. París, Ediciones UNESCO
- [4] Camacho, L., Zevallos, R. (2019) Siminchikkunarayku. Lingüística computacional para la revitalización y el poliglotismo. Hoja de Ruta. Fundación Siminchikkunarayku, Pontificia Universidad Católica del Perú.
- [5] Keselj, V. (2009). Speech and Language Processing Daniel Jurafsky and James H. Martin (Stanford University and University of Colorado at Boulder) Pearson Prentice Hall, 2009, xxxi+ 988 pp; hardbound, ISBN 978-0-13-187321-6.
- [6] Aronoff, Mark & Janie Rees-Miller (Eds.) (2003) *The Handbook of Linguistics*, Blackwell Publishers. ([ISBN 1-4051-0252-7](https://www.blackwellpublishing.com/9781405102527/)).
- [7] Unesco. (2003). Vitalidad y peligro de desaparición de las lenguas. Documento consultado en :http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/LVE_Spanish_EDITED%20FOR%20PUBLICATION.pdf
- [8] Chandía, A. (2012) Dingupeyem_alfa_v0.1: un prototipo de analizador morfológico para el mapudungun a través de transductores de estados finitos. . Master Interuniversitario de Ciencia Cognitiva y Lenguaje.
- [9] <https://news.un.org/es/story/2019/04/1454571>
- [10] Álvarez, A., Aranovic,R., Brown,R.,Carbonell, J., Fasola,C., Lavie, A., Levin, L., Font-Llitjos,A., Monson,C., Peterson,E., Probst,K. y Vega,R.M. Informe final proyecto Avenue/ Mapudungún: Desarrollo de herramientas informáticas para el mapudungún que se habla en Chile. (2005). Instituto de Tecnologías del Lenguaje, Universidad Carnegie Mellon.

- [11] Chandia, Andrés.(2013) Plan de investigación doctoral. Dunggupeyem: Analizador y generador morfológico a través de transductores de estados finitos y otras herramientas lingüístico-computacionales para el mapudungun. Universidad Pompeu Fabrá, Barcelona.
- [12] Marco Alvarado (2012). Sistema para el aprendizaje del mapudungun. Incluyendo características de reconocimiento de voz y bot conversacional. Pontificia Universidad Católica de Valparaíso.
- [13] CEDETi UC. Mapudungun mew.
<http://www.cedeti.cl/tecnologias-inclusivas/software-educativo/mapudungun-mew/>
- [14] Mager, M., Gutierrez-Vasques, X., Sierra, G. y Meza, I. (2018). Challenges of language technologies for the indigenous languages of the Americas. COLING 2018. Disponible en <https://github.com/pywirrarika/naki>
- [15] Ariadna Font Llitjós. (2005). Developing a Quechua-Spanish Machine Translation system.
- [16] Alex Rudnick. Towards Cross-Language Word Sense Disambiguation for Quechua.(2011). Proceedings of the Second Student Research Workshop associated with RANLP 2011 pages 133-138. Association for Computational Linguistics.
- [17] John Ortega, Krishnan Pillaipakkamnatt. Using morphemes for agglutinative languages like quechua and finnish to aid in low resources translation. 2018. Proceedings of the AMTA 2018 Workshop on Technologies for MT of Low Resource Languages (LoResMT 2018). Pages 1-11.Association for Machine Translation in the Americas.
- [18] Kann, K., Mager, M., Ruiz, I.V., & Schütze, H. (2018). Fortification of Neural Morphological Segmentation Models for Polysynthetic Minimal-Resource Languages. *NAACL-HLT*.
- [19] Arppe, A., Lachler, J., Trosterud, T., Antonsen, L., & Moshagen, S. N. (2016b). Basic language resource kits for endangered languages: A case study of Plains Cree. In C. Soria, L. Pretorius, T. Declerck, J. Mariani, K. Scannell, & E. Wandl-Vogt (Eds.), CCURL 2016—Collaboration and computing for under-resourced languages – towards an alliance for digital language diversity (LREC 2016 workshop), Portorož, Slovenia, 23 May 2016. European Language Resource Association.
- [20] Alicia Alexandra Alissi. Natural Language Processing and the Mohawk Language. Creating a finite state morphological parser of Mohawk formal nouns.2013. University of Limerick.

- [21] Epichan, A. M. y Oncevay, F. A.(2018). Language identification with scarce data: A case study from Peru. En *Information Management and Big Data: Fourth Annual International Symposium, SIMBig 2017*. Springer International Publishing.
- [22] Alva, C. A. y Oncevay, F. A.(2017). Spell-checking based on syllabification and character-level graphs for a Peruvian agglutinative language. En *EMNLP 2017 Workshop on Subword and Character Level Models in NLP, SCLeM 2017*. ACL Anthology.
- [23] Diiek Z. Hakkani-Tur, Kemal Oflazer, Gokhan Tur. Statistical morphological disambiguation for agglutinative languages. 2000. *COLING 2000 Volume 1: The 18th International Conference on Computational Linguistics*.
- [24] Ksenia Shalnova and Bruno Golénia. 2010. Weakly supervised morphology learning for agglutinating languages using small training sets. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING '10)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 976-983.
- [25] Loganathan Ramasamy, Zdeněk Žabokrtský, Sowmya Vajjala. The Study of Effect of Length in Morphological Segmentation of Agglutinative Languages. 2012. *Proceedings of the First Workshop on Multilingual Modeling*. Pages 18-24. Association for Computational Linguistics
- [26] Mathias Creutz and Krista Lagus. (2007). Unsupervised models for morpheme segmentation and morphology learning. *ACM Trans. Speech Lang. Process.* 4, 1, Article 3 (January 2007), 34 pages. DOI:<https://doi.org/10.1145/1187415.1187418>
- [27] McCarthy, A. D., Vylomova, E., Wu, S., Malaviya, C., Wolf-Sonkin, L., Nicolai, G., Kirov, C., Silfverberg, M., Mielke, S. J., Heinz, J., et al. (2019). The sigmorphon2019 shared task: Morphological analysis in context and cross-lingual transfer for inflection. In *Proceedings of the 16th Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 229–244.
- [28] Alvarado, G.(2020). *Gltopolítica de la Desigualdad: Ideologías del Mapudungun y el Español en Chile (2009–2019)*. CUNY Academic Works.
- [29] Coña, P.(2019)*Kuyfi mapuche chumgechi ñi azmogekeel egün. Editorial Genlol.* (edición completamente en Mapuzugun).
- [30] Manquilef, M,(1911). *Comentarios del Pueblo Araucano (La Faz Social)*.

- [31] Manquilef, M,(1911). Comentarios del Pueblo Araucano II (La Jimnasia Nacional).
- [32] Organizaciones mapuche convocan a novena marcha por el mapuzugun. 18 de Febrero de 2020.
<https://uatv.cl/2020/02/18/organizaciones-mapuche-convocan-a-novena-marcha-por-el-mapuzugun-e-nmarcada-en-el-dia-internacional-de-la-lengua-materna/>
- [33] Lengua y Emancipación Nacional, por Viktor Naqill Gomez
<https://www.mapuexpress.org/2016/02/17/lengua-y-emancipacion-nacional-por-viktor-naqill-gomez/>
- [34] Zúñiga, F. (2006), Mapudungun. El habla mapuche.. Santiago de Chile: Centro de Estudios Públicos.
- [35] Smeets, I. (2008) A Grammar of Mapuche Volumen 41 de Mouton grammar library, ISSN 0933-7636.
- [36] Creutz,M, Lagus K. (2007). Unsupervised models for morpheme segmentation and morphology learning. ACM Trans. Speech Lang. Process. 4, 1, Article 3 (January 2007), 34 pages. DOI:<https://doi.org/10.1145/1187415.1187418>
- [37] McCarthy, A et al.(2019) “The SIGMORPHON 2019 Shared Task: Morphological Analysis in Context and Cross-Lingual Transfer for Inflection.” Proceedings of the 16th Workshop on Computational Research in Phonetics, Phonology, and Morphology : n. pag. Crossref. Web.
- [38] Chiodi, F, Loncon Antileo, E. (1999). Crear nuevas palabras. Instituto de Estudios Indígenas. Universidad de La Frontera; CONADI.
- [39] Cañumil, Tulio Fernando. (2011). Estudio del idioma mapuche, Mapucezugun ñi gynezuam. Florencio Varela: Editorial Xalkan.
- [40] Trudell, B. (2014) The Multilingual Education (MLE) network phenomenon: advocacy and action for minoritized language communities. Multiling. Ed. 4, 17 .
<https://doi.org/10.1186/s13616-014-0017-y>
- [41] Loncon, E. (2010). Derechos educativos y lingüísticos de los pueblos indígenas de Chile. Revista de Inclusión Social y Equidad en la Educación Superior.

- [42] Catrileo, M. (2017). Diccionario Lingüístico Etnográfico de la Lengua Mapuche: Mapudungun-Español-English (Vol. 7). Ediciones Universidad Austral de Chile.
- [43] Zúñiga, F. (2007). Mapudunguwelaymi am?‘¿ Acaso ya no hablas mapudungun?’ Acerca del estado actual de la lengua mapuche. *Estudios públicos (Santiago)*, (105), 9-24.
- [44] Wittig, F. (2009). Desplazamiento y vigencia del mapudungún en Chile: un análisis desde el discurso reflexivo de los hablantes urbanos. *RLA. Revista de lingüística teórica y aplicada*, 47(2), 135-155.
- [45] Lagos, C. (2012). El mapudungun en Santiago de Chile: vitalidad y representaciones sociales en los mapuches urbanos. *RLA. Revista de lingüística teórica y aplicada*, 50(1), 161-184.
- [46] Olate Vinet, A., Alonqueo Boudon, P., & Caniguan Caniguan, J. (2013). Interactividad lingüística castellano/mapudungun de una comunidad rural bilingüe. *Alpha (Osorno)*, (37), 265-284.
- [47] Sadowsky, S., Painequeo, H., Salamanca, G., & Avelino, H. (2013). Mapudungun. *Journal of the International Phonetic Association*, 43(1), 87-96.
- [48] Croese, R. A. (2014). Tiempo verbal en mapudungun. *Lenguas y Literaturas Indoamericanas*, (1).
- [49] Araya, B. V., Castellví, M. T. C., & Fernández-Silva, S. (2019). Noun formation in Mapudungun: Productivity, genuineness and language planning. *Revista Signos*, 52(100), 615.
- [50] Alvarado, M. E. (2019). EL ‘NATIVO HABLANTISMO’ EN LA INVESTIGACIÓN DE LAS LENGUAS INDÍGENAS: EL CASO DEL MAPUDUNGUN EN CHILE. *Trabalhos em Linguística Aplicada*, 58(2), 795-825.
- [51] Hasler, F., Olate, A., & Soto, G. (2020). Origen y desarrollo del sistema evidencial del mapudungun. *Círculo de lingüística aplicada a la comunicación*, (81), 9-26.
- [52] DePalma, R. (2015). Learning a minoritized language in a majority language context: student agency and the creation of micro-immersion contexts. *International journal of bilingual education and bilingualism*, 18(4), 426-442.
- [53] Lamiquiz, V. (1975) *Lingüística española*. Sevilla: Publicaciones de la Universidad de Sevilla, pp. 433-434.

- [54] Albaigès, J. M. (2001) Diccionario de palabras afines – con explicación de su significado preciso. Madrid: Espasa-Calpe, p. 464.
- [55] Holzinger, A. (2005). Usability engineering methods for software developers. *Communications of the ACM*, 48(1), 71–74. doi:10.1145/1039539.1039541
- [56] Gillian R. Hayes. (2011). The Relationship of Action Research to Human-Computer Interaction. *ACM Transactions on Computer-Human Interaction*, 18(3), article 15.
- [57] Catrileo C, M. (1972). A tagmemic sketch of Mapuche grammar. *ETD Collection for University of Texas, El Paso*. AAIEP00811.
- [58] Pérez, C. (2015). Ziccionario Mapuzugun-Castellano. sin editorial.
- [59] Azümche: Nürüfkan mapuzugun mew <<https://azumche.cl/nurufkan/>>
- [60] Hämäläinen, M. (2019). UralicNLP: An NLP Library for Uralic Languages. *Journal of open source software*, 4(37), [1345]
- [61] Moeller, S., Kazeminejad, G., Cowell, A. and Hulden, M. (2018). A Neural Morphological Analyzer for Arapaho Verbs Learned from a Finite State Transducer](<https://aclanthology.org/W18-4802>)
- [62] The UQAILAUT Project. Inuktitut Computing.
<http://www.inuktitutcomputing.ca/Uqilaut/info.php>
- [63] Snoek, C., Thunder, D., Lóo, K., Arppe, A., Lachler, J., Moshagen, S., Trosterud, T. (2014). Modeling the Noun Morphology of Plains Cree](<https://aclanthology.org/W14-2205>)
- [64] Liu, Z., Jimerson, R., Prud'hommeaux, E. (2021). Morphological Segmentation for Seneca. 90-101. 10.18653/v1/2021.americanlp-1.10.

Anexos.

A. Código reglas de conversión entre grafemarios.

A.1 Unificado a Ragileo.

```
import re
def rulpawe(txt):
    txt=re.sub('ch','c',txt)      #Ch,ch   C,c
    txt=re.sub('d','z',txt)      #D,d     Z,z
    txt=re.sub('ng','d',txt)     #Ng,ng   G,g
    txt=re.sub('g','q',txt)     #G,g     Q,q
    txt=re.sub('d','g',txt)     #G,g     Q,q
    txt=re.sub('l_','b',txt)     #L_,l_   B,b
    txt=re.sub('ll','j',txt)     #Ll,ll   J,j
    txt=re.sub('n_','h',txt)     #N_,n_   H,h
    txt=re.sub('tr','x',txt)     #Tr,tr   X,x
    txt=re.sub('t_','t',txt)     #T_,t_   ---
    txt=re.sub('ü','v',txt)     #Ü,ü     V,v
    return txt
```

A.2 Unificado a Azümchefe.

```
import re
def rulpawe(txt):
    txt=re.sub('d','z',txt)      #D,d     Z,Z
    txt=re.sub('ng','d',txt)     #Ng,ng    G,g
    txt=re.sub('g','q',txt)     #G,g     Q,q
    txt=re.sub('d','g',txt)     #D,d     Z,Z
    txt=re.sub('l_','lh',txt)    #L_,l_    Lh,lh
    txt=re.sub('n_','nh',txt)    #N_,n_    Nh,nh
    txt=re.sub('tr','tx',txt)    #Tr,tr    Tx,tx
    txt=re.sub('t_','t\\',txt)   #T_,t_    T',t'
    return txt
```

A.3 Ragileo a Azümchefe.

```
import re
def rulpawe(txt):
    txt=re.sub(r'(?![sScCtT])h','nh',txt)  #H,h     Nh,nh
    txt=re.sub('(?![hH])c','ch',txt)      #C,c     Ch,ch
    txt=re.sub('b','lh',txt)               #B,b     Lh,lh
    txt=re.sub('j','ll',txt)               #J,j     Ll,ll
    txt=re.sub('(?![Tt])x','tx',txt)       #X,x     Tx,tx
    txt=re.sub('v','ü',txt)                #V,v     Ü,ü
    return txt
```


A.4 Ragileo a Unificado.

```
import re
def rulpawe(txt):
    txt=re.sub('(?![sct])h','n_',txt)    #N_,n_    H,h
    txt=re.sub('c(?![h])','ch',txt)    #Ch,ch    C,c
    txt=re.sub('z','d',txt)            #D,d    Z,z
    txt=re.sub('(?![n])g','ng',txt)    #Ng,ng    G,g
    txt=re.sub('q','g',txt)            #G,g    Q,q
    txt=re.sub('b','l_',txt)            #L_,l_    B,b
    txt=re.sub('j','ll',txt)            #Ll,ll    J,j
    txt=re.sub('(?![t])x','tr',txt)    #Tr,tr    X,x
    txt=re.sub('v','ü',txt)            #Ü,ü    V,v
    return txt
```

A.5 Azümchefe a Unificado.

```
import re
def rulpawe(txt):
    txt=re.sub('z','d',txt)            #D,d    Z,Z
    txt=re.sub('(?![Nn])g','ng',txt)    #Ng,g    G,g
    txt=re.sub('q','g',txt)            #G,g    Q,q
    txt=re.sub('lh','l_',txt)            #L_,l_    Lh,lh
    txt=re.sub('nh','n_',txt)            #N_,n_    Nh,nh
    txt=re.sub('tx','tr',txt)            #Tr,tr    Tx,tx
    txt=re.sub('t\\','t_',txt)            #T_,t_    T',t'
    return txt
```

A.6 Azümchefe a Ragileo.

```
import re
def rulpawe(txt):
    txt=re.sub('nh','h',txt)            #H,h    Nh,nh
    txt=re.sub('ch','c',txt)            #C,c    Ch,ch
    txt=re.sub('lh','b',txt)            #B,b    Lh,lh
    txt=re.sub('ll','j',txt)            #J,j    Ll,ll
    txt=re.sub('tx','x',txt)            #X,x    Tx,tx
    txt=re.sub('ü','v',txt)            #V,v    Ü,ü
    return txt
```

B. Orden de los sufijos.

Una de las fuentes utilizadas para ver el orden de los sufijos fue esta imagen, en la que se puede ver gráficamente el orden de los casilleros que llevan los sufijos del mapuzugun. Esta fue complementada con los documentos que se describen en las referencias.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
nie	kiyaw	ñma	ye	ge	fem	we	ka	ke	wye	la	a
kunu	küte	lel		uw	rume	me		pe		nu	fu
	küle	ñmu		mu	kantu	pa		tu		ki	afu
	meke				faluw	pu		rke			
						rpa					
						rpu					

Imagen tomada de <https://azumchefe.cl/nurufkan/>. [59]

C Comentarios de los usuarios.

C.1 Primera Prueba:

Comentarios al traducir palabras:

"hay diferencia entre las dos separaciones propuestas de li: (ella.. si yo)"

"Kümelkaley ñi wüzamel"

"no"

"Kümelkaley! Quizás se podría incluir el sufijo de tercera persona al incluir la separación?
mütxüm-fĩ-y"

"Kümelkaley"

"demasiadas opciones"

"Me parece que la estructura interna de esta palabra produce muchas ambigüedades en el análisis que hace el programa."

"no lo logre"

"En el ejemplo de "lhali" el programa era capaz de separar el morfema de modo hipotético (-l) del de tercera persona (-i), en este caso aparecen fusionados el de modo hipotético y de 2 persona singular (-lmi)."

"no lo logre / demasiadas opciones"

"Al igual que otro ejemplo (mogeltufiñ) existen demasiadas interpretaciones de la estructura interna de la palabra."

"sin comentario"

"Wüzamgeafuy chi "aye-ntu-"? Ti küzawpeyüm mu wüzamgefuy "ayen-tu", welu femgelu txokilafiñ."

"no"

"Kümelkaley"

"para mi nivel de conocimiento no se que significa"

"no me ayudo"

"igual no supe interpretar esa palabra "

"me aportó que lleva la particula "ke" de habitualidad, pero no se el significado de la partícula "el""

"El programa es amigable"

"Fue más difícil, me guié más que nada en el texto, no me queda muy claro si la traducción quedó correcta"

"Desconozco la palabra nentu"

"Es una palabra larga que está aglutinada, con la ayuda del programa entendí cada separación pero aún así logré traducirla"

"El programa, en parte, me ayuda a entender la palabra en su composición pero no así a entenderla o hecer la traducción de ésta"

"Agregué la traducción con el programa, con dudas de que esté correcto"

"aparecen dos separaciones iguales pero una dice "de el/ella/ellxs a tú "orden"". no comprendo qué significa."

"hay dos opciones de separar iguales"

"con los ejemplos de separación de esta palabra me di cuenta a qué se refiere con "orden""

"en el caso de tu = verbalizador, o (come en caso de comidas, tocar en caso de instrumentos, etc) se repiten dos veces las opciones -a-el / -ael"

"en la opción con elmu sale "de el/ella/ellxs a mi 2 "si te"". no debería salir "de el/ella/ellxs a ti 2 "si te""

"entregó en blanco la separación"

"ngelan"

"me pareció genial que separe las partículas, siento que las flechitas que sirven para ver el significado de cada componente de la palabra, debido a la ausencia de colores, hace que uno no se de cuenta inmediatamente que contienen la información jajs, osea me pasó que solo vi amu-a-fu-yu y no encontraba donde aparecía el significado de cada partícula, entonces creo que falta algo o más visual o explicito que diga "desliza hacia abajo", kimlan chumgechi piafun feychi zugu.. aparte "afu", entiendo, cobra otro significado al estar unido y quizás podría considerarse como una sola partícula al momento de separarla amu-afu-yu"

"al comienzo no me di cuenta de que había que apretar el botón triángulo para el desglose del significado "

"no me apareció nada respecto a esta palabra, luego lo intenté en unificado escribiendo "ilelнгеkey" y ahí si me funcionó... ahora cuando me funcionó me impresionó la cantidad de formas de entender esa misma palabra, en este sentido, creo que hace que sea una herramienta muy útil para quienes estudiamos mapuzugun, solo que para una persona que recién se está acercando va a ser chocante, pero de todas formas está útil"

"no mostró la "separación""

"kvmeley"

"fue de mucha ayuda, solo me queda la duda si la definición de "ge" como "ojo" está bien que se explicita? osea creo que como partícula, por la ubicación en la que se encuentra dentro de la palabra, esa definición de "ojo" solo podría confundir, obviamente lo digo bajo mis conocimientos"

"la partícula ge tiene muchos significantes por lo que es necesario un contexto para "traducir" la palabra"

"gelay zugu"

"me gustó mucho esta forma de separar la palabra, me ayudó bastante y me enseñó cosas que no comprendía tan bien respecto a la partícula "nie" y tampoco conocía el verbo "lelinien""

"en vez de decir VER dice VTR "

"dumi-ñma-le-le"

"genial, asombroso lo funcional que es la separación de las partículas y conocer lo que pasa con algunos verbos que componen por ellas (no sé como se dice eso jajs, pero es genial notar la diferencia entre zumiñma-le-le y zumiñ-ma-le-le)"

"entristecerá-deprimir-oscurerá "

"gelay zugu"

"la verdad el hecho de mostrar "fu" como algo que ya no corre, me hizo pensar en que la traducción de niekellefuy es: ya no suele tener... pero siento que eso sería niewekelay, entonces creo que eso podría llegar a confundir. De todas formas entiendo la necesidad de expresar el fu como algo que dejó de ser, pero quizás algún ejemplo sería bueno, "niefun: había tenido""

"no se me ocurrió una verbo habitual que ya no corre..."

"gelay zugu"

"entiendo que sería bueno incorporar otras definiciones para la partícula tu, además de: "verbalizador, o (come en caso de comidas, tocar en caso de instrumentos, etc)", pues creo que significa "volver a hacer algo" o en otros casos expresa cambio. Y en este caso indicaba aquello a mi parecer"

"se me hace confuso la indicación de " lo que queda o (con negación)""

"gelay zugu"

"no conocía "kel", solo conocía "kil" y me quedó un poco la duda con o también mezclado con condicional (si es que), es decir, en este caso no está negando? se ser así sería "si es que me compra algo"... ahí tengo la duda"

"me parece confusa la indicación "algo que ya no corre (FU pero seguido de E se convierte en FE)""

"gelay zugu"

"me sirvió mucho, pero siento que pudo ser repetitiva la separación en este caso. De todas formas me ayudó a entender mejor el significado a mi parecer"

"tirar y rechazar me parecen significados super diferentes"

""o como diciendo "dicen que"(porque no es testigo directo o lo sabe 100%)": esta definición la he leído en varios libros de gramática, pero ningún mapuzugufe me la ha corroborado, más bien dicen que se usa para "darle color" a lo que uno está relatando"

"La definición de -ñma- es un poco confusa."

"La definición de -ge- es un poco confusa."

Comentarios al traducir el texto:

"se casi nada de mapuzugun, así que me habría tomado mucho tiempo"

"rume kvmengey tami programa peñi, tutelkaley anay! kvme kvzaw peñi, lemoría!"

"en "no" falta una L en la definición pues dice "negación para lo condiciona y otros modos," usé la herramienta en algunas palabras que no conocía o que tenía duda y fue de ayuda, solo que la palabra kakewme no me apareció"

"A medida que voy ocupando el programa me familiarizo un poco más, aunque me cuesta entender la forma de explicar las palabras que están aglutinadas"

"Con la primera parte que ví, la herramienta me pareció espectacular, de mucha utilidad tanto para traducir una palabra suelta como para aprender. Gracias. Felicitaciones por sacar adelante esta brillante idea."

"No utilicé la herramienta para traducir el texto. "

"Me gustó mucho la herramienta para entender el texto. Me gusta que da la definición de la palabra y las partículas, sobre todo en las formas con etew, fiel, etc. Me ayuda a ver rápido quien hace a quien; cuando lo hago con una tabla que tengo me demoro en que "me caiga la teja" pero como acá está dicho directo "de tal a tal, o se usa de 3as a cualquiera, etc" es rápido. Me gusta que muestre todas las combinaciones aunque hayan algunas que son imposibles porque me sirve para entender que hay algunos sentidos que no pueden ser si las partículas fueran esas (por ej. kvpal-el-a-en vs kvpa-le-la-en). Me gusta que es muy rápida y que diga que detecta este u otro grafemario. En aspecto al principio confunde un poco pero con un par de usos se van entendiendo las cosas, como que se pueden apretar las palabras que entrega separadas o ciertos fraseos que hay en las definiciones (por ej de tu a yo "orden"). Quizás sería útil un "glosario" que aclare cosas como vi, vtr, de x a y, etc. Rume kvme y tami hemvl wvzalkawe peñi."

"no tiene la opción de traducir el texto completo entonces toma demasiado tiempo ir palabra por palabra, de hecho no es palabra por palabra, sino una amplia variedad de posibles significantes"

Lo que más gustó a los usuarios :

"que da una buena idea de los sentidos que puede tener una palabra"

"Como separa las palabras por partículas."

"La separación y las alternativas ayudan a entender muy rápido el significado."

"me gustó que mostrase distintas formas de separar una palabra en sus significados, dejando la posibilidad de indagar más allá en cada expresión que se escucha. A veces uno entiende la raíz y no comprende el uso o separación de las partículas (sobre todo cuando se juntan las vocales), en este sentido la herramienta es de utilidad, a veces pasa al revés que no entendemos la raíz verbal pero existen más diccionarios donde indagar, en este sentido, la herramienta innova respecto a las partículas y el uso de diversos grafemarios. Dos grandes aciertos que permiten facilitar el aprendizaje del mapuzugun"

"no existe otra de este tipo"

"Permite hacer un estudio mas acabado de cada palabra, eso si que no sirve para principiantes, para una persona que no sabe nada de la gramática tanto del mapudungun como del castellano se le hará muy difícil, sin embargo la lógica del aglutinamiento de las palabras puede ser internalizada con el uso constante de esta herramienta."

"Que permite comprender los diferentes sentidos que tienen las palabras al combinarse las partículas."

"que funciona, sirve (una vez conocida su manera de operar."

"Que sea capaz de hacer una predicción del grafemario y que esté en mapuzugun."

Lo que menos gustó a los usuarios:

"que algunas veces repite mucho. "

"No entrega traducción al wigkazugun."

"El aspecto."

"Creo que evidentemente sería bueno mejorar aspectos estéticos con fines didácticos. Además, a futuro, sería genial pensar en audios. Lo digo tmb porque participo en un taller de mapuzugun con una persona ciega y me imagino que si ella sigue aprendieron le sería de utilidad algo así."

Además, algunas partículas como afu no se entendieron juntas y creo que hubiese sido bueno, y otras como "tu" sólo contaban con un significado, en una de las palabras que busqué. En este sentido, si bien es un desafío lograr hacer esto. Ya han desarrollado una herramienta genial que tiene un buen futuro, newentu!"

"la gráfica"

"No siempre entendí (como este: no/siempre) la explicación del desglose de una palabra, eso me confundió y se me hizo complejo una vez que quise realizar la traducción del texto mas largo. Aunque, creo que una vez que mas familiarizado se avance con menos dificultad. "

"Me parece que la información está muy desordenada, sobretodo en las palabras con muchos significados. "

"que el recuadro de ingreso de palabras a separar es poco vistoso; muy chico en comparación con el resto de la página. Pero esto es un pelo de la cola, con la utilidad que presta/brinda."

"La parte visual, eché de menos uso de colores para distinguir los morfemas. La doble columna español / mapuzugun, quizás podría ser una columna donde el mapuzugun esté arriba y el español abajo?"

Comentarios generales:

"por ahora nada"

"nada peñi, esta muy buena su herramienta. Le deseo mucho éxito en su tesis."

"Me gustó mucho. Es muy útil y fácil de usar. Destaca que te deja listo el trabajo de buscar en un diccionario la raíz o palabra y qué partículas hay para que uno interprete. Yo cuando leo y no sé palabras uso el diccionario en WordTheme de azumchefe.cl y me demoro mucho entre ingresar todas las opciones que podrían ser la raíz hasta dar con la que más se adecue. Además, como la herramienta lee cualquier grafemario se salva más tiempo aún, por ej, en mi diccionario está todo en unificado y el libro que leo en azümchefe, entonces a veces busco y el diccionario me dice que no existe el término pero pasa que yo lo ingresé en el grafemario incorrecto (ej. buscar lhafkenh cuando en la app de diccionario necesita que escriba l'afken'; o darme cuenta que escribí mal la l y no la n y buscar lhafken' o cosas por el estilo) y no me di cuenta y seguí de largo."

"rume mañum müten, newentuleaymi chacha"

"lamento no haber podido hacer la parte del texto, pero de verdad encuentro que toma mucho tiempo, y ahora no lo tengo, lo siento. "

"Primero que todo felicitar, es un tremendo trabajo, vaya que kudaw! y si bien es cierto mi mapudungun es básico creo que me ayudó bastante más que usar un diccionario."

"Muy seguramente servirá mejor a aquellos que manejan la gramática castellana. Por otro lado, testeé la herramienta una sola vez desconociendo la topografía de todo el programa, o sea entré a "capella" por lo tanto uno da vueltas más largas intentando descubrir de qué se trata todo esto, cómo lo abordo. En definitiva, debo decir que soy en esto una usuaria básica, distinto es ver cómo sería si lo usara por mucho tiempo o que fuera una herramienta para un curso donde se estaría usando de manera continua."

"Se podría ofrecer una tanto traducción literal como adaptada para facilitar la comprensión del sentido de las palabras. "

"Ya lo dije todo. Vuelvo a felicitarte."

"Me costó deducir que se puede desplegar la traducción de los morfemas apretando la flecha negra y lamento no haberme dado cuenta antes, pero me parece muy buena herramienta ahora que noté ese detalle."

C.2 Segunda Prueba:

Comentarios al traducir la palabra “*pemurpayafuyu*”:

"No la reconoce, por lo que pide volver a intentar con otra palabra "

"No se encuentra"

"No había escuchado esta composición. Chaltumay"

"Newe mülekelan contexto ñi nentual feychi -mu kiñe nütamkan mu! Welu fewla küme azümniefiñ"

"difícil por el desconocimiento de las partículas"

"Me pasariN a ver"

"No tengo comentarios."

"Kimlan, no pude abrir el programa "

"Ayüfin chumngechi ñi wüdamfiel feychi troyke nemül. Küme femi."

"La explicacion de yu me confundio un poco "

"Me sirvió para darme cuenta de mi error, había confundido la transición -mu- con la articula -ñmu-"

"Esta palabra está escrita de diferente manera a la anterior, pero las partículas son las mismas, se entiende igual."

"Es una palabra confusa, me cuesta interpretarla"

"No se entiende cuando no se domina la lengua "

"Solo si conoces la lengua puedes hacer uso del traductor"

"entiendo cada parte por si sola pero no logro comprender su traducción"

"me cuesta traducir esta palabra desde el programa"

"los colores, en la parte del rojo no se distingue bien, además algunas traducciones ocupan mucho espacio y se pierde un poco el ir palabra a palabra "

"Kuzawgei!"

""accion mientras se viene" is confusing me :)"

Comentarios al traducir la palabra “*Elukelafimu*”:

"Podría agregar ejemplos, puesto que el sentido de la acción no queda del todo claro "

"En las separaciones el -mu final hace que salga duplicada cada separación porque un mu es "Ustedes (2 personas)" y el otro "que uds 2 lo hagan""

"bien especificado el significado de las partículas"

"Me confundí en esta, me quedan las dudas de cuando las terminaciones pueden ser ambiguas. Por ejemplo ("kel"/"a") o ("ke"/ "la")"

"Kimlan, no pude abrir la página me sale “unauthorized”"

"Küme wüdamnofilu feychi nemül, trokiniefiñ."

"elu ke la fimu"

"No conocía la partícula -kel-, aunque pareciera ser lo mismo que -kil-, pero no conocía el modo imperativo combinado con futuro. Hay una forma de separar la palabra que no me parece correcta porque toma la partícula -mu como imperativo y a la vez la partícula -la- como negación, siendo que en imperativo se debe negar usar -kil-. Fey müten pepi feypiafun"

"La separación -kel- no la tenía en consideración. Pero excelente forma de explicar."

"Entiendo la separación, pero no estoy segura de que significa específicamente"

"La segmentación era clara. Sin embargo, propone 2 traducciones posibles, y me parece que sólo una es posible ya que la negación del imperativo es -kil y no -kel "

"No es muy claro para las personas que son principiantes o que buscan traducir ,es poco amigable y complejo."

"Muy complejo de encontrar la traducción si eres principiante,"

"No entendí como obtener la traducción a partir de la separación"

"Las otras separaciones pueden generar confusiones para we chillkatufe"

"No entiendo dónde aparece el significado en sí de la palabra :("

"Esta difícil, pero creo que me sirvió el programa."

"No se lo dieron, entregaron ustedes dos"

Comentarios al traducir la palabra “*Kujinerkeeiñmu*”:

"Faltaría un ejemplo donde se use Welulkan ñi trürngenukunual rke/ke!"

"la descripción del ne dice "tamién""

"me ayudo bastante el programa aquí No conocía la partícula "ne" y se me hizo un poco complejo comprenderla, también que "kullin" estuviera con "j" me confundió al principio"

"Küme femi."

"Me ayudo a percatarme de la partícula -ne-"

"El programa habla de él/ella/ellos(as) pero la termina es -mu que se refiere a ustedes dos"

"Nunca había visto esa palabra «kuji» me confundio su traduccion, no estaba clara decia vtr "

"Hay que ser hablante para entender"

"Es complejo cuando no se maneja el idioma usar la herramienta de traducción queda la duda sobre el significado de la palabra hay typos en la traduccion"

"Aquí estoy complicado..."

"the "kuji" label says: "vi & vtr pagar". What is vi and vtr? Also, in the label for "ne", the "tamién" should be "también", maybe??"

"Kisu zunguan mapuzugun kiñe futa nemüll(grafemario), recomendable utilizar sólo un grafemario"

Comentarios al traducir la frase “*Pichikalu iñche , amukefun chillkatuwe ruka mew , fewla chillkatuwekelan .*”:

"chillkatuwekelan, esa palabra no arrojó traducción "

"Buena aplicación, estoy aprendiendo harto"

""1-Cuando una misma particula tiene mas de un posible significado dentro de una misma forma de separar la palabra podría aparecer una acepción abajo de la otra sin volver a repetir la palabra diseccionada entremedio, para hacer mas fácil la lectura."

"2- Al principio dudé de la utilidad de que el programa separe mew en me-w, por ejemplo, al no tener la terminación -ün (mewün), ya que me parece casi imposible que la palabra mew pueda significar hacerse excremento sin tener una terminación verbal, y además esas separaciones podría prestarse para interpretaciones ziley-móricas de muchas palabras si el usuario no tiene criterio. Sin embargo, luego pensé que en algún texto o al escuchar de oído alguna frase, uno

podría llegar a encontrarse algo como ""mew egün"" o ""mew ymi"" y en esos casos podría llegar a ser útil.""

"Al final ,entendí más como se usaba la herramienta, creo que el que va a traducir no maneja la lengua entonces no se entienden las instrucciones ,al principio puse la frase entera ,luego puse por separado las palabras y tenía más opciones. Igual encuentro poco amigable la herramienta un hablante la entiende pero una persona no hablante que quiera traducir es complejo."

"Hay que familiarizarse más para poder opinar"

"la palabra chillkatuwe tiene muchos significados que tal vez no ayuden a llegar al significado final de "escuela" o "colegio". Para las palabras con mayor aglutinación se requiere tener un manejo en mapuzugun para no confundirse."

"this is hard!"

"Newe ayülañ feychi "verbo puro" zungun, chumngelu azümniefuy "puro"? "

"es un tanto mas difícil entender la traducción en las frases"

"En "chillkatuwe ruka", la explicación de "tu" no me quedo tan clara."

"Ayüñ tufachi düngu: ", Tufachi nhemül pelañ, ka kiñe tukulfe"."

"No ocupe la herramienta para traducirla."

"Me.gustaría que saliera altiro el significado y además la separación"

"Es muy interesante, pero aún no me siento seguro de mis traducciones, aunque podría ser cosa de práctica y de habituarse."

"Escribe aquí si es que tienes comentarios de la herramienta y de su funcionamiento para apoyarte a traducir la frase."

"Si uno separa las palabras que no son verbos muestra opciones de separación como que fueran verbos por ej. cijkatuwe, ruka y fewla. Iñce también muestra una separación extraña iñ+ce. "

Lo que más gustó a los usuarios:

"Ayuda mucho a realizar la separación, como persona que está aprendiendo aún me cuesta hacer ese procesamiento de las diferentes partículas, en especial, en palabras más extensas. Ayuda muchísimo a comprender el sentido de las partículas."

"Lo fácil y sencillo de usar"

"Es útil especialmente en los estadios iniciales e intermedios de aprendizaje. Parece ser muy eficiente con palabras complejas, con varios morfemas, como las consultadas. No la probé con otras palabras. Además, la incorporación de colores la hace amigable."

"Tengo que familiarizarme más con ella para que me pueda gustar siento que la práctica de su uso me permitirá darme cuenta de su valor y el trabajo que implica crear una herramienta de este estilo."

"Permite ejercitar las conjugaciones verbales, y quizás puede ayudar a traducir con más práctica."

"Que las explicaciones de las separaciones son claras."

"Pronunciacion"

"No la entendí"

"El mismo hecho de las separaciones de sufijos propios del idioma "

"que da varias opciones y uno puede elegir la que mas le tinca porque no soy hablante "

"nada"

"La dinámica desafiante, hace pensar, "

"Explica parte por parte de la palabra"

"Que aprendes la conjugacion"

"Las diferentes propuestas para desglosar las palabras"

"Es rapida, facil de entender y me gusta que no de una traduccion resuelta y requiera un esfuerzo del chillkatufe para comprender."

"Muy dificil. "

"Que permite entender gráficamente la composición de las palabras en mapuzungun."

"el buscar es fácil"

"Que permitiera ingresar palabras de mucha aglutinación"

"La utilidad para aprender la gramática"

"La explicación detallada de cada partícula."

"las definiciones de las partículas y la posibilidad de escribir en distintos grafemarios, así como la posibilidad de poner frases para traducir"

"Doy ayüfiñ ñi kúme adngen fachi düngu. Ka kúme wüdamkey ñi wirintükuelchi ñemül iñche."

"la explicación de las partículas que pueden ayudar a aprender su uso y contexto "

"La descomposición de las partículas y su significado por separado "

"La posibilidad de conocer algo de manera simple que es muy difícil de encontrar."

"Que todo estuviera en mapuzungun!"

Lo que menos gustó a los usuarios:

"que se como ver la traduccion"

"Puede ser que algunas definiciones a veces no fueran tan sencillas de entender, sin embargo, traspasar el lenguaje gramatical a uno más sencillo es muy desafiante. Quizá sería positivo contar con el acceso a un PDF que pudiera uno comprender más sobre el uso de las partículas."

"En el caso de la parte estética, recomendaría usar colores que fueran más parecidos o armónicos, recomendaría esta página para crear paletas de colores <https://htmlcolorcodes.com/es/selector-de-color/>"

"Mm. Quizás debiera tener palabras en español, a fin de que hablantes iniciales aprendan mejor y no se pierdan"

"Algunas traducciones me parecieron tan informales y breves que pueden inducir a error. Por ejemplo, traduce el morfema -fu como "algo que ya no corre". Alguien desinformado puede entender eso como la acción de correr. De manera similar, se traduce -rpa como "acción mientras se viene", lo que es poco claro. Quizás podrían agregarse ejemplos sencillos para clarificar."

"Otro asunto que vislumbro como posible problema es el que la herramienta entrega significados morfemáticos, mientras que en especialmente en algunos casos los significados son el resultado de combinaciones de morfemas, como en la combinación -we -la (ya no) o -pe -la (que literalmente morfema a morfema se entiende como acción reciente negativa, pero se suele usar para afirmar algo, con un sentido modal)."

"Que para un no hablante que quiere traducir ,no hay instrucciones de uso,por instinto se va aprendiendo su uso ."

"Nada, felicitaciones!"

"Me pasa que siento que sin conocimiento hay algunas opciones de separación que uno podría confundirse. Me pasó con palabras que mostraban una separación donde había -kel- y -a- y creo, pero no estoy seguro, que esa combinación no se da, pero como no lo sé la consideré como una posible traducción de una "orden negativa a futuro" o algo así. Pero en el caso de que no se de esa combinación no tengo la habilidad para descartarla por mi cuenta. Pero eso pasa en palabras solas, si uno va trabajando un audio o texto el contexto puede ayudar a descartar también (por ejemplo en la frase yo escogí una separación de cijkatuwekelan que tenía que ver con la oración (ya no estudio), que fuera una orden no tenía sentido y la descarté)."

"Error escrita de palabras"

"No tiene instrucciones de uso para los principiantes "

"nada"

"Es digital, y actualizada"

"No entrega (o no lo vi) el significado final de la palabra, o ejemplos en dónde usarla, o cómo usarla"

"Que no da la traducción inmediata, deja mucho a la interpretación personal"

"No se si que no me gusto, pero me parece que el hecho de que si o si haya que escribir gramaticalmente correcto puede traer dificultades. Por ejemplo cuando hay 2 e. Por ej, busque poyeyu y poyeeyu. "

"L a escritura "

"Las manera de representar el tipo de palabra ("s", "vtr", "vi") no son claras para una persona que no las ha visto antes."

"la separación de las palabras, colores utilizados, a ratos se duplicaban, y el espacio muchas veces hace difícil de leer, no es muy fácil de utilizar para alguien que está recién aprendiendo"

"Que no ofrezca una traducción, aunque aquello sería muy difícil, imagino"

"Un poco tediosa de leer"

"Se repetía dos veces cada traducción, creo que no es necesario."

"Posibilidades de traducción erróneas, por ejemplo, en chillkatuwe, dar como posibilidad "we" como "ya no" (al acompañarse de negación)."

"Fey wüdamfilu kiñeke ñemül, tripay wedake troyke ñemül. Feychi -mu (2° Indicativo mo), welu doy pepi kûmekûnungelu, trokiniefñ."

"quizás el detalle en la explicación de las partículas, ya que puede ser algo muy interpretativo "

"Que no entregue ejemplos "

"A veces como que se demoraba un poco, no sé si es mi conexión tal vez."

"Los botones de mapuzugun y español podrían ser un poquitín más visibles. "

"Para un usuario que no entiende mucho (lamentablemente) puede ser complicado. Pondría más traducciones a futuro."

Comentarios generales:

Comentarios finales, inquietudes o dudas. ¿Qué puedes decirnos sobre la herramienta que no haya estado en alguna otra pregunta? "

"Considerar que no todas las personas manejan la cantidad de partículas que se utilizan en el mapudugun, por lo que sería interesante adjuntar una sección de glosario (quizás se tiene, pero desde el celular es un poco difícil recorrer la página en general). "

""Está muy bacán! Nunca había visto algo similar así de novedoso!"

"Tal vez muy a futuro se podrían agregar audios con las traducciones para conocer la pronunciación. "

"Muchas gracias por compartir este gran avance :D""

"Kûme zugu peñi!! Mañumeyu ta herramienta. Lemorria"

"Pensé que sería más fácil "

""En tema de colores, cuando hay suficientes partículas en la separación se muestran en colores diferentes y algunos son oscuros y dan poco contraste a las letras por lo que cuesta un poco leerlos bien."

"También podría ser útil fijar el color a ciertas partículas para ir asociando significados visualmente, por ejemplo -a- siempre amarillo, -la- siempre naranja, etc. así cuando la persona separe varias palabras luego de un tiempo al solo ver los colores puede rápidamente obtener el sentido. ""

"muy buena herramienta, y mas para mi que intento aprender partículas, que a mi parecer son muchas y dependen del contexto "

"Solo tengo conocimiento de palabras, pero no de oraciones"

"La escritura,la pronunciacion de las palabra de como se escriben,,"

"Excelente herramienta, creo que ayudará mucho a entender la gramática del mapudungun y que puede ser una excelente herramienta para quienes estamos en un nivel intermedio, ya que permite revisar palabras o frases de mayor complejidad. Igualmente, ayuda a entender mejor las características y lógica del mapudungun. Felicito el enorme trabajo, lamngen. Müna küme küdaw zewmaimi, küpe tripape tüfachi dungu, lamngen."

"No pude utilizar el programa, no me abrió "

"Küme femimi/femimün! Ayüfiñ fachi düngu, fey mew ta iñ pu che niey kümeke rakiduam, eyimi/eymu/eymün küme longko nielu, newentuleaymi/ymu/ymün!"

"nada"

"Me parece una idea buenisima. Seria de mucha ayuda para lxs chillkatufe a quienes les ayuda mucho mas la explicacion gramatical."

"Que diccionario(s) usa la herramienta?"

"Excelente programa, claramente necesita pulirse un poco, pero es una excelente iniciativa y se agradece la intención de expandir el mapuzugun a las demás personas!"

"Van a sacar una app ? Seria más rápido para usarla"

"Muy buena iniciativa. En mi caso no sé nada sobre mapuzungun y herramientas como esta son muy atractivas y útiles para comenzar a aprenderlo."

"En las 2 últimas preguntas con escala, la pregunta no clarifica si 1 o 5 es la mejor nota. Yo consideré que 5 es la mejor nota."

"Valoro el trabajo realizado para poder tener esta herramienta y valoro que la sometan a evaluación así podrán tomar en cuenta las opiniones y si es necesario mejorar los detalles.Felicitaciones por la iniciativa."

""La herramienta se ve bastante útil. Este es mi primer acercamiento al mapuzungun, tengo 0 conocimiento, pero gracias a la herramienta pude intentar traducir. "

"La dificultad para traducir es alta, pero creo que a mí se me hizo particularmente difícil por no tener una base en el idioma. "

"¡Excelente iniciativa!"

"Siento que no me fue útil, pero porque no entendí como usarla"

"Me. Costó mucho "

"Creo que tendría que habituarme más a su uso para recomendarla y para entender su potencial y limitaciones."

"Recomiendo q sea más atractiva, con imágenes orientadoras. No estamos en un nivel ideal, y debe ser algo más empático con las y los usuarias/usuarios"