



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

---

**FACULTAD DE INGENIERÍA**

**Creación de un lenguaje  
artificial a partir de  
programación genética**

**TESIS**

Que para obtener el título de

**Ingeniero en Computación**

**P R E S E N T A**

Jorge Iván Negrete Rodríguez

**DIRECTORA DE TESIS**

Dra. Katya Rodríguez Vázquez



Ciudad Universitaria, Cd. Mx., 2026



**PROTESTA UNIVERSITARIA DE INTEGRIDAD Y  
HONESTIDAD ACADÉMICA Y PROFESIONAL  
(Titulación con trabajo escrito)**



De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado CREACION DE UN LENGUAJE ARTIFICIAL A PARTIR DE PROGRAMACION GENETICA que presenté para obtener el título de INGENIERO EN COMPUTACIÓN es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi Entidad Académica, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de titulación.

---

**JORGE IVAN NEQUETE RODRIGUEZ**  
Número de cuenta: 307244011

*A mi familia que quiero tato a'tal cielo:  
mi abuelita, mi mamá, mi tía y mi hermano.*

*A mis amigos por aguantar mi depresión  
e impulsarme siempre hacia adelante.*

# Agradecimientos

A mi familia: a mi mamá, Magdalena; a mi tía, Ethel; a mi abuelita Socorro; y a mi hermano, Mauricio; por estar siempre presente en mi vida. Por apoyarme a pesar de mis momentos difíciles. Abuelita, te quiero mucho.

A Ricardo Cesar por su amistad, la fraternidad que hemos construido y por siempre estar ahí. Porque sin su apoyo esta tesis no estaría terminada. En verdad, gracias.

A Ricardo Enríquez por ser esa luz cálida en mi vida, por quedarnos en esta simulación.

A Daniel M. Olivera por haber cambiado mi visión del mundo con su taller de literatura y cine de mal gusto.

Al M. en I. Jesús Edmundo Ruiz Medina por su amistad durante la carrera. En otra vida volveremos jugar Maple.

A la Doctora Katya Rodríguez Vázquez por su apoyo incondicional en este largo proceso y por darme la oportunidad de concretar este proyecto.

A todos los que me han dado su apoyo. Gracias.

I read somewhere the trick is to keep breathing...

—Janice Galloway, *The Trick Is to Keep Breathing*

# Resumen

En este trabajo se desarrolla un programa bio-inspirado para su aplicación en el proceso de creación de un lenguaje artificial. El motor está centrado en que la evolución genética encuentra muchas similitudes con la evolución de un lenguaje. El lenguaje no es un fenómeno estático, evoluciona tal como los seres vivos en la naturaleza. Además, los lenguajes artificiales son proyectos de lingüistas y de aficionados que toman demasiado tiempo en realizarse, e incluso, nunca llegan a finalizarse. Es aquí donde esta investigación propone un auxiliar para la creación de un lenguaje artificial mediante la aplicación del cómputo evolutivo. Específicamente se programó un generador de palabras a partir de programación genética, que se expandió con elitismo y reemplazo, y se utilizó un *treebank* para definir las reglas de producción de la gramática del nuevo lenguaje. Estos resultados, las palabras generadas y la gramática, facilitan la creación de un lenguaje artificial en cuestión de días. Reducen esfuerzo al momento de imaginar palabras y en la definición de reglas de producción para el nuevo lenguaje. Sin embargo, este programa funciona mejor siendo una herramienta de apoyo, más que una solución. El proceso creativo del individuo es esencial para darle vida a un proyecto como la creación de un lenguaje artificial.

# Abstract

A bioinspired program was developed for this thesis to be applied in the creation process of an artificial language. What drove me to it was the similarity between genetic evolution and the evolution of language. Language is not a static phenomenon; it evolves just as living organisms in nature. Since artificial languages are projects of the linguists and non-professionals, and they take a long time to achieve, generally, they never get to be finished. Aiming to resolve this problem, this thesis offers a tool to help in the creation of artificial languages through the application of the evolutionary computation. Specifically, a word generator was programmed based on genetic programming, expanded with elitism and replacement, and a *treebank* was used for the definition of the production rules for the new language's grammar. These results, the generated words and the new grammar, do ease the creation of a new artificial language, it even reduces the time to create one in a few days. It reduces the effort at the time of imagining words, and in the definition of the production rules for the new grammar. Nevertheless, this program works better as an assisting tool, rather than the sole solution to the problem. The creative process of individuals is essential to give life to a project like it is one of creating an artificial language.

# Índice General

Índice Figuras.....	VIII
Índice de Tablas .....	X
1 Introducción.....	1
1.1 Estructura de la tesis .....	3
2 Marco teórico.....	4
2.1 Cómputo evolutivo.....	4
2.1.1 Algoritmos genéticos .....	6
2.1.2 Programación genética.....	11
2.2 Aspecto lingüístico.....	13
2.2.1 Contexto.....	13
2.2.2 Tipología lingüística .....	14
2.2.2.1 Tipología sintáctica.....	14
2.2.2.2 Tipología morfológica .....	14
2.2.3 Gramática.....	16
2.2.3.1 Gramática generativa .....	19
2.2.4 Fonotáctica.....	21
2.2.5 Léxico .....	25
2.2.5.1 Afiliación genética.....	25
2.2.5.2 Identidad e integración.....	25
2.2.5.3 Patrimonio y cultura.....	25
2.2.5.4 Préstamos .....	26
2.2.5.5 Dialectos .....	26
2.2.5.6 Cambios de significado.....	27
2.2.6 Morfosintaxis .....	28
2.2.7 Semántica.....	29
2.3 Lenguaje artificial .....	29
2.3.1 Lenguajes <i>a priori</i> y <i>a posteriori</i> .....	29
2.3.2 <i>Auxlangs</i> .....	30
2.3.2.1 Esperanto.....	30
2.3.2.2 Interlingua.....	31

2.3.2.3	Lengua de señas .....	32
2.3.3	Engelang .....	35
2.3.3.1	<i>Loglan</i> .....	35
2.3.3.2	Toki pona .....	36
2.3.4	artlangs .....	37
2.3.4.1	Lenguajes élficos .....	38
2.3.4.2	<i>klíngon</i> .....	40
2.3.4.3	trunic .....	40
2.3.5	Lista de palabras .....	41
3	Desarrollo .....	43
3.1	Diseño del algoritmo.....	43
3.1.1	Planteamiento del problema.....	43
3.1.2	Hipótesis y soluciones esperadas .....	44
3.2	Primera fase de desarrollo.....	44
3.3	Segunda fase de desarrollo.....	56
3.3.1	Desarrollo de gramática .....	58
3.3.2	Reducción de las reglas de producción final.....	73
4	Daino gazezuko .....	77
5	Conclusiones y futuros trabajos.....	86
6	Bibliografía.....	89
7	Apéndice.....	93
7.1	Instrucciones para usar el código.....	93

# Índice Figuras

Figura 2.1 Flujo convencional de un algoritmo genético (Koza, 1992) .....	6
Figura 2.2 Individuo en programación genética, un árbol donde los nodos internos son funciones y los nodos terminales valores .....	12
Figura 2.3 Cruza de dos individuos; en el primer individuo intercambia la primera rama y el segundo individuo la segunda rama.....	12
Figura 2.4 Mutación del individuo, en la que una rama del individuo muta en una función .....	13
Figura 2.5 Descomposición de una oración para el análisis de sus componentes, en la que se analiza un Sintagma Nominal .....	17
Figura 2.6 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Adjetival.....	17
Figura 2.7 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Preposicional.....	17
Figura 2.8 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Verbal.....	18
Figura 2.9 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Adverbial.....	18
Figura 2.10 Alfabeto fonético internacional, traducido al español. Se desglosa el sistema de notación estándar utilizado por lingüistas para representar los sonidos del habla de cualquier idioma de manera precisa y unívoca.....	23
Figura 2.11 Alfabeto fonético internacional actualizado. Se desglosa el sistema de notación estándar utilizado por lingüistas para representar los sonidos del habla de cualquier idioma de manera precisa y unívoca.....	24
Figura 2.12 Entrada del diccionario de Lengua de Señas Mexicana, descripción de la seña, apoyo visual y análisis lingüístico. ....	34

Figura 2.13 Entrada del diccionario de Lengua de Señas Mexicana, descripción de la seña, apoyo visual y análisis lingüístico. ....	34
Figura 2.14 Treinta y seis símbolos de los ciento veinte símbolos que componen el lenguaje Toki Pona.....	37
Figura 2.15 Guía para escribir en inglés utilizando el sistema de escritura Tengwar, creado por J.R.R. Tolkien .....	38
Figura 2.16 Leyenda de Trunic, comparativa entre Trunic e inglés. ....	40
Figura 2.17 Guía para escribir en inglés utilizando el sistema de escritura Trunic .....	41
Figura 3.1 Componentes de un individuo: cromosoma, aptitud, punto de cruce, progenitor uno y progenitor dos .....	45
Figura 3.2 Apoyo visual de la reproducción de dos individuos, los dos nuevos individuos y de la posible cruce y mutación que ocurre a partir de los parámetros de cada uno de los individuos. .	51
Figura 3.3 Ejemplo de las funciones de evaluación dentro del programa para las diferentes longitudes, en este caso, longitud tres y longitud cinco.....	52
Figura 3.4 Análisis de una oración a partir de las reglas gramaticales de la nueva gramática .....	57

# Índice de Tablas

Tabla 1.1 Correspondencia entre conceptos de evolución natural y sus contrapartes computacionales.....	7
Tabla 2.1 Lista de palabra swadesh de cien palabras.....	42
Tabla 3.1 Referencia de probabilidad acumulada para la selección de caracteres en el lenguaje	46
Tabla 3.2 Referencia de valor de importancia para la evaluación .....	47
Tabla 3.3 Referencia de valor para casos especiales en los que un tipo de letra específico necesite ser seleccionada .....	47
Tabla 3.4 Reglas fonotácticas para la evaluación de palabras .....	49
Tabla 3.5 Resultado de la ejecución del programa, un conjunto de palabras que han sido evaluadas .....	54
Tabla 3.6 Asignación de las palabras generadas a la lista swadesh.....	56
Tabla 3.7 Análisis de cincuenta oraciones en diferentes idiomas a partir de las reglas de la nueva gramática, y las reglas de producción generadas a partir de los análisis. ....	72
Tabla 3.8 Tabla final de reglas de producción después de una reducción .....	74
Tabla 3.9 Ejemplos de oraciones en el nuevo lenguaje, después de haber obtenido las reglas de producción y el conjunto de palabras.....	76
Tabla 4.1 Traducción de las frases utilizadas en el cuento .....	84

# 1 Introducción

En los últimos años, los lingüistas han intentado destruir la barrera del lenguaje entre individuos de diferentes países. Para ello, Okrent (2009) habla de que se han creado lenguajes como el *esperanto*, o Van Steenberghe (2020) del *budinos*, para facilitar la interacción entre un grupo de comunidades con dificultades para comunicarse en su vida diaria. Además, Adams (2011) confirma que la creación de lenguajes es algo que recientemente ha tomado fuerza en el entretenimiento. Ahora se crean lenguajes para criaturas que existen únicamente en historias que llegan a ser cuentos, novelas, series de televisión, películas o videojuegos.

Sin embargo, este proceso puede tomar mucho tiempo, desde un par de meses hasta incluso años. Por ejemplo, L.L. Zamenhof, el creador del *esperanto*, trabajó por lo menos siete años antes de realizar su primera publicación; J. R. R. Tolkien desarrolló el lenguaje de los elfos en un año, mientras estudiaba, pero trabajó durante toda su vida en la evolución y afinación del lenguaje; Gina H. Amelio, trabajó en el *Enkárík*, para Frankelda, durante un año. Este tiempo se usa para definir el alfabeto, las fonotácticas, el cómo va a sonar cada sílaba, la gramática, el contexto histórico en el que se desarrolla este lenguaje, qué sí y qué no se puede comunicar, derivando declinaciones, mapeando campos semánticos, revisando el avance del lenguaje, etc.

Fue un taller de Daniel M. Olivera, en la Facultad de Química, sobre creación de lenguajes artificiales, o *conglangs*, que despertó mi interés por crear uno de autoría propia para alguna obra, ya fuera un videojuego o una novela. En ese momento también realizaba mi servicio social en el IIMAS bajo la supervisión de la Dra. Katya Rodríguez, aprendiendo sobre programación genética. El punto de partida fue la intersección de estas dos experiencias: el estudiar sobre lenguajes artificiales y programación genética. Esta intersección permitió identificar lo laborioso del proceso

de creación de un lenguaje como un sistema que podía ser automatizado sin perder su esencia de lenguaje 'natural'. ¿Y no es esta una de las capacidades que otorga la ingeniería en computación? La automatización de procesos complejos de otras disciplinas. Así, se reconoció que el proceso evolutivo simulado en los algoritmos genéticos podría ser análogo a la evolución histórica de un lenguaje natural.

Si bien existen trabajos que aplican la computación a la lingüística, sus esfuerzos se han concentrado principalmente en el análisis más que en la síntesis. Las investigaciones han buscado analizar la gramática de los lenguajes, generar diálogo simple, generar un subconjunto de gramáticas o gramáticas con características específicas, o visualizar la evolución de ciertos lenguajes. Aunque valiosas, abordan problemáticas de asistencia y análisis, pero no el desafío de la creación generativa de un lenguaje artificial.

Por ello, el objetivo de este trabajo es un programa bio-inspirado configurable que permita automatizar la creación de un lenguaje artificial: el programa generará el vocabulario que se adapte a las necesidades de un lingüista o aficionado: qué consonantes y vocales el lenguaje va a utilizar; la estructura de las palabras; qué prioridad en sonidos y combinación de ellos va a utilizar. Esto dará como resultado un lenguaje prototipo: un grupo de palabras que funcionará de vocabulario. Para la gramática se utilizarán *treebanks* para su definición. Debe cumplir con características específicas de sonido, fonotáctica y gramática. Por último, he decidido utilizar el lenguaje generado en un cuento como ejemplo.

## 1.1 Estructura de la tesis

En el Capítulo 2 se describen las bases teóricas tanto de cómputo evolutivo, del aspecto lingüístico y de los lenguajes artificiales. El Capítulo 3 describe el desarrollo de la solución, dividido en dos partes. La primera parte dedicada a la generación del vocabulario y la segunda parte para la generación de la gramática. El Capítulo 4 es un cuento de autoría propia en el que se utiliza el lenguaje generado. Y, finalmente, en el Capítulo 5 se exponen las conclusiones obtenidas durante este proyecto e ideas para trabajos futuros.

# 2 Marco teórico

## 2.1 Cómputo evolutivo

A lo largo del tiempo las especies han cambiado de una forma u otra. Los rasgos que definen a los individuos se han transformado un sinnúmero de veces a través de generaciones.

Luchando por sobrevivir y reproducirse, dichos individuos heredan sus genes a una nueva generación, provocando cambios que, por muy pequeños que sean, son decisivos en la evolución de las especies. El proceso evolutivo requiere una constante adaptación a las nuevas circunstancias, obligándoles a competir por recursos en el ambiente en el que se desarrollan, para sobrevivir o para decidir el liderazgo del clan. Sólo aquellos que sobresalgan en su desempeño podrán heredar sus genes a la siguiente generación. Así, a través de la selección natural, los individuos son cada vez más aptos, mejores en lo que hacen.

Tomando en cuenta estas premisas sobre la selección natural, entre 1950 y 1970, Lawrence Fogel, John Holland, Ingo Rechenberg y Hans-Paul Schwefel, investigadores en el campo de la inteligencia artificial, fueron precursores en el campo del cómputo evolutivo. Generaron teorías y algoritmos para la automatización de solución de problemas tomando particularidades de los seres vivos, la evolución y la supervivencia del más apto. Tres teorías destacan entre las demás: ‘programación evolutiva’ (Fogel, 1966), ‘algoritmos genéticos’ (Holland, 1967) y ‘estrategias evolutivas’ (Rechenberg, 1965), ya que más tarde se convertirían en los tres paradigmas centrales del cómputo evolutivo.

Estos paradigmas permitieron atacar problemas de los que anteriormente era difícil obtener soluciones. Específicamente, problemas en los que no existe una solución única sino un grupo considerable de soluciones, en donde a medida que el conjunto de soluciones en cuestión crece,

más complicado y tardado es obtener una solución o conjuntos de soluciones óptimas con métodos convencionales. La computación evolutiva permite entonces resolver ciertos problemas que los métodos directos (ej. que la función a evaluar no tenga derivada) y heurísticos (ej. al evaluar de forma aleatoria se pueden perder de vista mejores soluciones) no son capaces de resolver.

El objetivo fundamental del cómputo evolutivo es encontrar una buena solución a problemas que no tienen un único resultado, a partir de un grupo de posibles soluciones generado de forma aleatoria. Su pseudocódigo, basado en el comportamiento evolutivo de las especies, es simple:

- Generar un grupo de soluciones de forma aleatoria.
- Elegir de entre el grupo pares de soluciones para intercambiar características.
- Las soluciones generadas a partir del intercambio tendrán probabilidad de mutar sus características.
- Evaluar a las nuevas soluciones y decidir si cumplen con los criterios de solución.
- Si no se ha encontrado un individuo ideal (una solución al problema), se reinicia el proceso a partir del segundo paso.

En la práctica, la primera generación es generada de forma aleatoria sin considerar si las soluciones tienen características deseadas. Cada uno de los individuos, una posible solución, se evalúa de acuerdo con pruebas de aptitud y se califica qué tan buena es su combinación de características. Después, dependiendo del método de selección, se intercambian características entre individuos para generar nuevos individuos que contengan características mixtas. Este nuevo grupo pasará por el mismo proceso de evaluación para descartar soluciones que no se acerquen a las características deseadas. Dependiendo del filtro establecido por el usuario, se decide si es necesario volver a intercambiar características y crear un nuevo grupo o si en el grupo ya existe una solución que se acerca a la solución deseada.

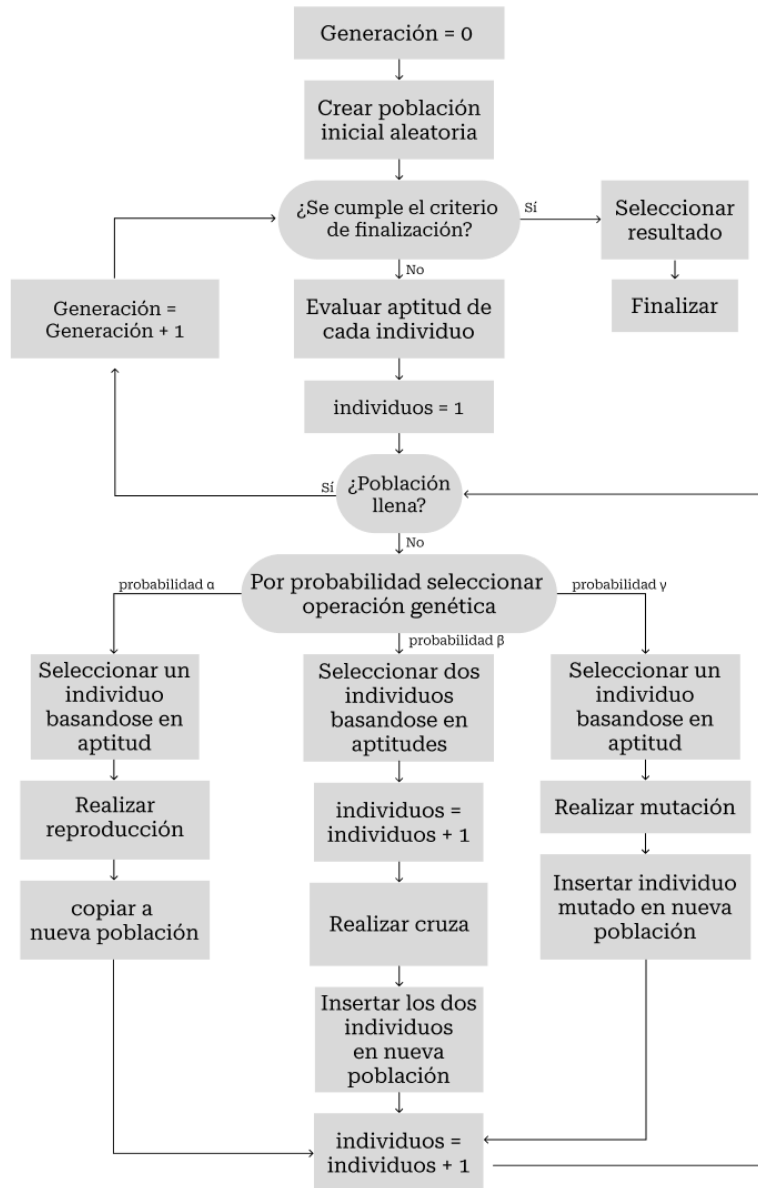


Figura 2.1 Flujo convencional de un algoritmo genético (Koza, 1992)

### 2.1.1 Algoritmos genéticos

A partir del trabajo de John Holland (1975), *Adaptation in Natural and Artificial Systems*, aparece la rama de algoritmos evolutivos. En su trabajo introduce varias investigaciones y experimentos que se venían trabajando desde diez años antes, en el área de la evolución simulada. Esta visión

fue uno de los primeros avances concretos en adaptar los sistemas de supervivencia de los seres vivos a la resolución de problemas.

Así, el vocabulario utilizado en la teoría de la evolución se convirtió en algo aplicable en el ámbito de la computación. Un gen se convirtió en un carácter, un cromosoma en una cadena de caracteres, el genotipo de un individuo en la estructura de la cadena y un alelo en el valor de uno de los caracteres. Junto con esta nueva terminología se introdujeron también métodos de selección, operadores de cruce, mutación y elitismo.

En su libro, Holland (1975) ya no habla de un solo individuo o posible solución a un problema, sino de una población de individuos que sufrirá cambios a lo largo de generaciones. Se reproducirán, tendrán descendencia y morirán, hasta que en alguna generación un individuo de la población tenga las características ideales que se buscan.

Holland (1975) propuso simbolizar a un individuo como una cadena de bits. Cada bit o fragmento de la cadena representaría un parámetro para una posible solución, que más tarde sería decodificada. Al decodificar todos los componentes de un individuo es posible saber si éste es una solución capaz de resolver satisfactoriamente el problema o no.

Una vez que se conoce la capacidad de cada uno de los individuos de la población, dependiendo del método de selección, se cruzan los individuos para obtener descendencia que contenga una cadena más variada, con características de ambos individuos padres seleccionados para cruzar.

### Terminología

Naturaleza	Algoritmo Genético
Gen	Carácter   Bit   Elemento
Alelo	Valor
Cromosoma	Cadena
Genotipo	Estructura

Tabla 1.1 Correspondencia entre conceptos de evolución natural y sus contrapartes computacionales.

## **Individuo**

Un individuo está compuesto por el conjunto de información considerado como una solución al problema. Fogel et al. (1997) argumentan que del problema y el tipo de solución que requiera, la estructura del individuo se adapta lo más posible. Desde la forma básica para representar un individuo, una cadena de bits, hasta una compleja lista de elementos de diferentes naturalezas.

A partir de la forma y las características, cada individuo tendrá cierta aptitud y competirá contra los demás individuos en la población.

## **Operadores genéticos**

Al igual que ocurren cambios en el ADN de los seres vivos, en el momento en el que se cruzan los individuos existen operadores genéticos que introducen diversidad genética en la población modificándolos en diferentes formas. Estas modificaciones permiten poner a prueba un mayor número de soluciones. Sin embargo, el uso de estos operadores debe ser restringido, su abuso podría perjudicar la eficiencia de la búsqueda al perder individuos con características deseadas por mutar genes aleatoriamente. Los operadores más utilizados son la cruce entre individuos y la mutación de los genes.

## **Cruza**

El objetivo de cruzar dos individuos es tener la posibilidad de poder obtener las mejores características de ambos. La cruce genera  $n$  número de fragmentos de dos individuos denominados padres. Fogel et al. (1997) describen que estos fragmentos se intercambian a partir de cierta probabilidad para generar nuevos individuos en la nueva generación. Como resultado, los nuevos individuos comparten características de los individuos padre de la generación anterior.

Holland (1975) propuso, sin embargo, realizar la cruce en un solo punto, es decir, sustituir todos los bits desde un punto elegido hasta el final de la cadena, al cual llamó cruce en un punto. Este

punto de cruce casi siempre es tomado de forma aleatoria en la mayoría de los casos, permitiendo recombinar tanto pequeños como grandes fragmentos. Se debe notar que la cruce puede tener un alto grado de impacto pues tiene la oportunidad de modificar grandes partes de los individuos. Sin embargo, esto no significa que la cruce entre individuos sea algo negativo, al contrario, su existencia añade una variabilidad diferente al de la mutación. Dicha variabilidad resulta del uso de genes que ya existían en la generación anterior, encaminando así, a una mejora de la población.

### **Mutación**

Fogel et al. (1997) describen la mutación como el cambio de una sola característica de un individuo. Dependiendo del genotipo (la estructura del individuo), la modificación puede ser de un bit, un carácter o un número. Siempre modificando un solo gen, ya que cada uno de los elementos del individuo es independiente. Un caso de mutación mínimo es aquél en el que sólo muta un elemento y en el caso de mutación máximo es aquél en el que muta la cadena entera.

### **Función de desempeño**

Una función de desempeño o función objetivo es por la cual se juzgan las aptitudes de los individuos. De acuerdo con qué características tengan o no, obtienen una calificación. Fogel et al. (1997) plantean que deben diseñarse de manera particular para cada problema. Ya sean particularidades negativas o positivas, se tiene que especificar qué tanto beneficia o disminuye la aptitud del individuo.

### **Aptitud**

La aptitud se define como la fuerza que impulsa a un individuo en la selección natural. Ésta se valora dependiendo de la capacidad de adaptación del individuo para que pueda reproducirse y dejar descendencia. Este concepto se aplica íntegramente a los algoritmos genéticos. Una vez

evaluados los individuos se les otorga cierto peso, dependiendo de si contienen características deseadas o no. Se puede medir de forma explícita o implícita, aunque la forma más común de hacerlo es de forma explícita, al asignar un valor escalar al individuo, de acuerdo con Fogel et al (1997).

### **Selección por ruleta**

Fogel et al. (1997) plantean que la selección por ruleta se encarga de seleccionar qué par de individuos serán utilizados como padres en el método de cruce. En este tipo de selección, la probabilidad de que un individuo sea seleccionado depende de su propia aptitud con respecto a los demás en la población, es decir, entre mejores características tenga con respecto a los demás individuos en la población, mayor será la probabilidad de ser seleccionado y, de igual forma, si no contiene características deseadas, el individuo tendrá muy poca probabilidad de ser seleccionado.

### **Selección por torneo**

Koza (1992) habla sobre la selección por torneo, una manera muy similar a los torneos que se realizaban en la antigua Grecia, en las que dos combatientes peleaban por sobrevivir. La adaptación de este torneo a los algoritmos genéticos selecciona dos individuos al azar y compara sus aptitudes. El individuo que tenga mejores características tendrá la posibilidad de cruzarse.

### **Elitismo**

El elitismo, aunque no es utilizado tanto como una forma de seleccionar individuos para la siguiente generación, funciona muy bien como una herramienta auxiliar que garantiza que no se pierdan los individuos más aptos a lo largo de las generaciones de acuerdo con Fogel et al. (1997).

## 2.1.2 Programación genética

En 1992, John R. Koza expande el paradigma de los algoritmos genéticos e introduce la programación genética en sus libros de *Genetic programming: On the programming of computers by means of natural selection*. Entre las innovaciones que introduce están: la introducción de una forma jerárquica en forma de árbol para representar las soluciones, permitiendo solucionar problemas ambiguos. Generando de forma implícita una variabilidad dinámica, el tamaño y la forma de cada programa impactan directamente la aptitud de dicho programa. Un programa muy grande o un programa con subrutinas demasiado extensas podrían tardar en terminar en una cantidad excesiva de tiempo.

Esencialmente, la modificación al algoritmo genético evolutivo es en el genotipo, considerando a un árbol un individuo. Para facilitar la representación del árbol se utilizan expresiones S, cadenas de caracteres, secuencias de símbolos, elementos terminales y no terminales, bajo reglas que manejan el orden de los símbolos. Una máquina de estados finitos en donde todas las operaciones están cerradas y no contenga errores en la secuencia. Entre las funciones que se podrían utilizar están las funciones lógicas: AND, OR o NOT; funciones aritméticas, trigonométricas, exponenciales; funciones definidas por el propio usuario, que modifiquen resultados a operaciones existentes; o macros, un número de operaciones en cierto orden. Puede también contener valores TRUE, FALSE, variables numéricas, operadores comparativos condicionales u operadores de ramificación condicional, siempre y cuando no exceda el tamaño límite restringido por el usuario.

### **Individuo**

El individuo en programación genética consiste en una expresión S, un árbol en el que cada nodo interno puede ser una función y nodo terminal, nodo hoja, contiene un valor a ser utilizado, de acuerdo con Koza (1992). Al momento en que se inserta una función, con ella se inserta el número

de hojas necesarias para realizarla, y así hasta que las últimas hojas insertadas sean todas elementos terminales.

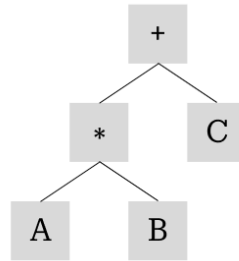


Figura 2.2 Individuo en programación genética, un árbol donde los nodos internos son funciones y los nodos terminales valores

### Cruza

Koza (1992), para programación genética define que la cruce puede resultar un poco agresiva, por el hecho de que la formación de los programas es a partir de árboles y es indispensable que no se pierda dicha estructura. Por lo que, al seleccionar el elemento a cruzar, se seleccionará a partir de una raíz hasta el final de todas las hojas que contenga. Puede llevar a cruzar desde una sola hoja hasta un árbol completo. A su vez, los elementos que se cruzarán no son necesariamente del mismo tamaño, ambas selecciones se realizan de forma individual.

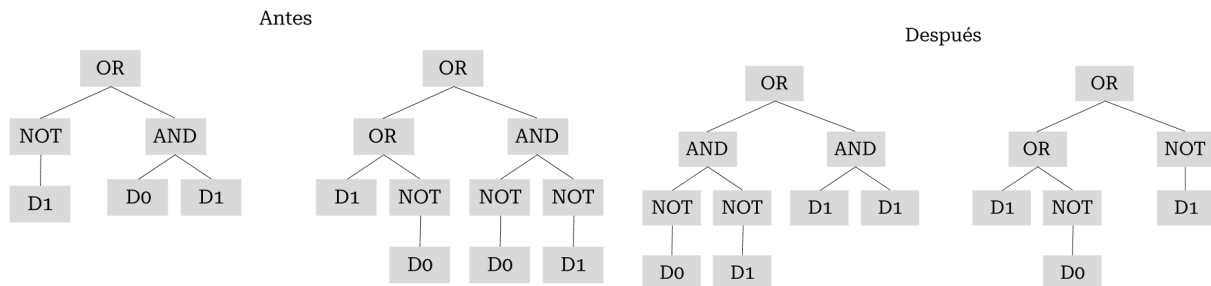


Figura 2.3 Cruza de dos individuos; en el primer individuo intercambia la primera rama y el segundo individuo la segunda rama

### Mutación

Para programación genética, Koza (1992) define que la mutación no es en un solo carácter. Ya que se maneja a partir de árboles, para facilitar su evaluación, es necesario mutar desde una raíz, y mutar el sub-árbol completamente. En caso de no hacerlo, podría corromperse el individuo. Esta

mutación es un nuevo árbol, tiene la probabilidad de ser una simple hoja o un árbol de varios niveles, siempre y cuando se respete el nivel límite.

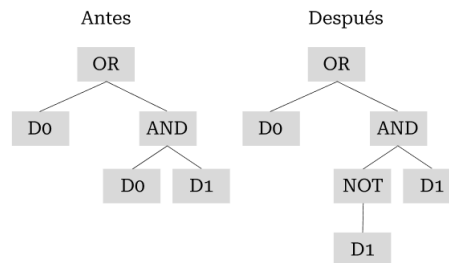


Figura 2.4 Mutación del individuo, en la que una rama del individuo muta en una función

## 2.2 Aspecto lingüístico

### 2.2.1 Contexto

El lenguaje es uno de los factores que en el desarrollo humano ha permitido al hombre desenvolverse y comunicarse con otros, transmitir emociones, deseos e información. Por ello, en toda lengua se hace un estudio de su evolución histórica y su estructura.

La biolingüística es el análisis a partir de la biología y la evolución del lenguaje. En los últimos años uno de los representantes más importantes ha sido Noam Chomsky. Él ha planteado la posibilidad de que en el lenguaje los elementos se unen de forma externa, es decir, constituyen un componente de una estructura más grande. De ahí que Chomsky vea la similitud entre un lenguaje y un programa. Por tanto, se crea el “Programa minimalista”, una rama de investigación en la que se concentraría todo lo relacionado con el lenguaje, la programación y su evolución artificial.

## 2.2.2 Tipología lingüística

### 2.2.2.1 Tipología sintáctica

Song (2010) describe que la tipología lingüística se encarga de clasificar la estructura base de los lenguajes, es decir, las similitudes y los rasgos lingüísticos que se encuentran en la sintaxis, el orden de las palabras en una oración, y el orden morfológico, la estructura de las palabras.

Desde el punto de vista de la sintaxis, los lenguajes tienen similitudes en su estructura con un limitado número de combinaciones en cuanto al orden de los elementos en la oración Sujeto (S), Objeto (O) y Verbo. (V).

- El orden SOV es utilizado en idiomas como el japonés, coreano y latín.
- El orden SVO es utilizado en idiomas como el inglés, alemán y chino.
- El orden VSO es utilizado en idiomas como el galés, mixteco y árabe.
- El orden VOS es utilizado en idiomas menos comunes como el fiyiano y malgache
- El orden OSV es utilizado en menos idiomas, popularmente lo utiliza el personaje Yoda.

### 2.2.2.2 Tipología morfológica

El intentar tipificar los lenguajes del mundo, Greenberg (1974) describe, ha ocurrido desde ya un par de siglos. Es en 1808, con las definiciones propuestas por los hermanos Friedrich von Schlegel y August von Schlegel, cuando se consolidaron dos clasificaciones de los lenguajes:

- Lenguas analíticas: en las que palabras raíz se acompañan de preposiciones, postposiciones, partículas y modificadores, utilizando muy raramente afijos. En ellas las palabras comúnmente sólo tienen un significado y no es posible derivar unas de otras. Entre los lenguajes que utilizan esta clasificación está el noruego y sueco. Por ejemplo, en noruego: qué – hva; hva heter du? – ¿cómo te llamas?

- **Lenguas aisladas:** en las que las palabras en su totalidad consisten sólo de palabras monoformáticas. Entre los lenguajes que utilizan esta clasificación está el chino. Por ejemplo, en donde cada uno de los kanji representa una idea, 子為誰 – [el otro] – ser - ¿quién?
- **Lenguas sintéticas:** en las que palabras raíz se acompañan de prefijos y sufijos para cambiar el significado o emoción.
  - **Lenguas aglutinantes:** En esta clasificación la formación de palabras se da cuando a una palabra se le agregan sufijos y prefijos establecidos para cambiar el significado o emoción. Entre los lenguajes que utilizan esta clasificación está el japonés, náhuatl, finlandés, esperanto y turco, entre otras. Por ejemplo, en el japonés: Aka – rojo; Akairo - color rojo; akambou - infante; akaji – déficit; akagane – cobre. O en el purépecha donde los plurales se forman con el sufijo -echa, -icha o -cha.
  - **Lenguas flexivas:** en donde las palabras están formadas por una palabra raíz, que pueden reducirse a uno o más elementos para modificar su significado. Entre los lenguajes que utilizan esta clasificación están el ruso, alemán, español, latín y portugués. Por ejemplo, en el español la raíz ‘gat-’ puede formar las palabras gato, gata, gatos, gatas, gatitos.

Esta categorización ayuda de cierta forma como base para la investigación de los idiomas y ha funcionado por gran cantidad de años. En los últimos cincuenta años, los lingüistas han llegado a un consenso en el que consideran al lenguaje como un ente generativo. En el que consideran que la naturaleza del lenguaje es algo que los seres humanos internalizamos y utilizamos como mejor nos convenga. No es una máquina que puede ser definida en su totalidad, pero sí quizá en gran parte.

### 2.2.3 Gramática

La gramática es el estudio de las reglas que rigen a los lenguajes a partir de la estructura fonética-fonológica, sintáctica-morfológica, léxica-semántica y pragmática de las palabras y oraciones.

En términos generales, LLorach (1994) define que las reglas de producción son:

< **Oración** > → < *S. Nominal* > < *S. Verbal* >

< **Sintagma Nominal** > → (*Determinante*) < *S. Nominal* > < *S. Adjetival* >  
< *S. Preposicional* > < *Oración* >

< **Sintagma Adjetival** > → (*Intensificador*) < *S. Nominal* > < *S. Adjetival* >  
< *S. Preposicional* > < *Oración* >

< **Sintagma Preposicional** > → < *Preposición* > < *S. Nominal* >

< **Sintagma Verbal** > → < *Verbo* > < *S. Nominal* > < *S. Adjetival* > < *S. Preposicional* >  
< *Oración* > < *S. Adverbial* >

< **Sintagma Adverbial** > → (*Intensificador*) < *Adverbio* >

Una oración es construida a partir de su estructura bipolar; las palabras que forman una oración no se agrupan libremente, sino que siguen un orden en sus constituyentes, el sintagma nominal y el sintagma verbal.

La estructura del sintagma nominal consiste en cinco elementos: un determinante; un sustantivo que puede ser común o propio, contable o no contable; un sintagma adjetival; un sintagma preposicional y/o una oración adicional.

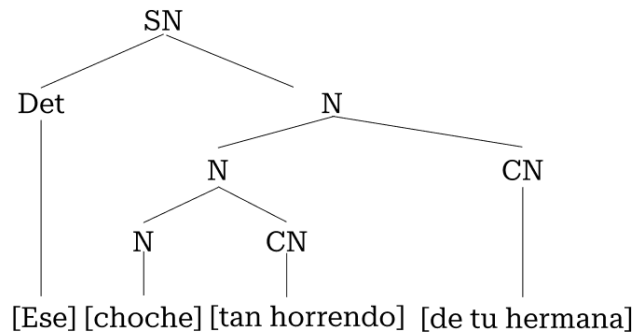


Figura 2.5 Descomposición de una oración para el análisis de sus componentes, en la que se analiza un Sintagma Nominal

El sintagma adjetival consta de un intensificador o adverbio, un núcleo y posibles complementos.

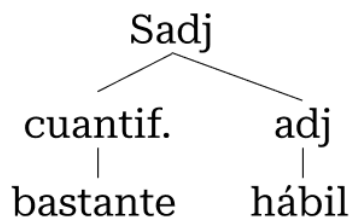


Figura 2.6 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Adjetival

El sintagma preposicional se estructura a partir de una preposición y un sintagma nominal.

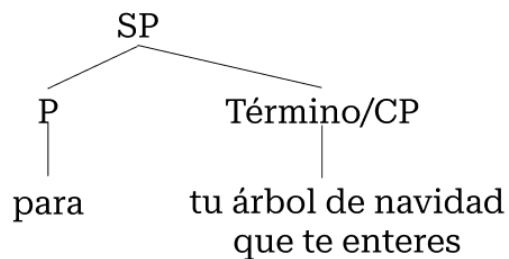


Figura 2.7 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Preposicional

El sintagma verbal está constituido de un verbo y diferentes complementos, ya sea un sintagma nominal, un sintagma adjetival, un sintagma preposicional, un sintagma adverbial u otra oración.

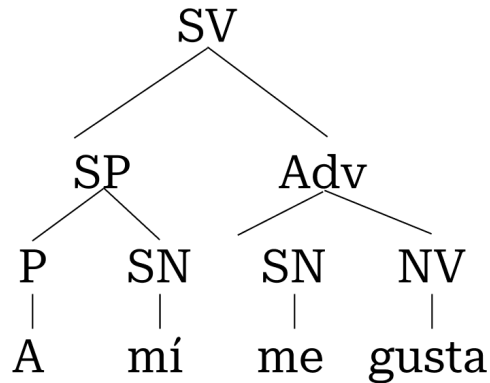


Figura 2.8 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Verbal

El sintagma adverbial consiste en un intensificador y un adverbio.

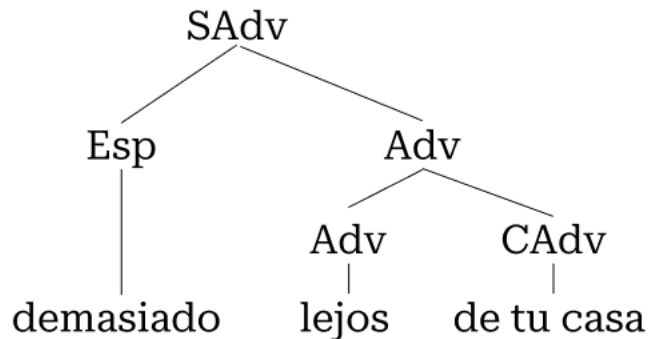


Figura 2.9 Descomposición de una oración para el análisis de sus componentes, en la que se analiza el Sintagma Adverbial

Por último, los demostrativos, que LLorach (1994) desarrolla, son palabras que sitúan de forma más precisa al sintagma nominal que delimita el espacio-tiempo al que se hace referencia. En la mayoría de los lenguajes se tienen dos o tres niveles de distinción, en casos muy raros existen cuatro. Esto influye en la forma en que las personas se comunican cuando desean precisar información a partir del demostrativo. En el español existen tres niveles: este, ese y aquel, cada uno con sus respectivas flexiones de género y número. En inglés sólo hay dos niveles: this y that. En japonés son tres niveles: *これ, それ, あれ*.

### **2.2.3.1 Gramática generativa**

La capacidad básica del ser humano para comunicarse ha permitido a los individuos transmitir información a partir de cualquier forma de expresión. Sin importar si entre el emisor y el receptor existe una forma de expresión en común, oral, escrita o corporal, los individuos lograrán comunicarse. A esta capacidad intrínseca de intercambio de información, Aguilar (2004) en un artículo sobre Chomsky, define como “competencia lingüística”. Una aptitud que por sí sola moldea nuestra realidad a través de simbolizar y evocar imágenes cognitivas, que nos permite evolucionar nuestro propio lenguaje al reconocer palabras nuevas y asimilarlas a un campo semántico de posibilidades y no a un objeto o acción concreta.

Sin embargo, la necesidad de transmitir ideas claras obligó a los grupos a crear normas que regularan las formas de expresión, otorgando al lenguaje una forma concreta y coherente. Definidas en conjunto como gramática, estas normas regulan las características sintáctico-morfológicas, el contexto y estructura de las palabras en la oración, sin descartar totalmente la influencia que tiene en la formación fonética y léxica de las palabras.

Si la gramática, Chomsky pone en consideración, es un conjunto de reglas finito que permite generar todas y cada una de las manifestaciones lingüísticas ya existentes de un lenguaje, entonces se podrían formar conjuntos de reglas que puedan generar sub-lenguajes o incluso nuevos lenguajes.

Sin embargo, el análisis para encontrar todas y cada una de las características que definen a la gramática más simple de un lenguaje data, desde el final de la década de 1940 a la fecha, no se han encontrado y lo más probable es que no se encuentren. La permisividad del lenguaje, aquella que permite crear oraciones mal estructuradas, puede crear problemas de contradicción entre las reglas de un conjunto.

De esto que se haya considerado un acercamiento biológico para comenzar un análisis más detallado de los lenguajes. En él, Chomsky (2005) considera que el lenguaje tiene un crecimiento a partir de herencia, experiencia y principios en la competencia lingüística que son un poco más ambiguos y no son tan fácil de precisar. Justo como cualquier ser vivo, tiene la capacidad de crecer y evolucionar a lo largo de los años, agregando palabras o modificándolas de acuerdo con su contexto.

De ello que lingüistas consideren hallar conjuntos de reglas que constituyan la teoría para la gramática para cada uno de los lenguajes que se conocen. Sin embargo, para elaborar una teoría lingüística se toma en cuenta:

- El conjunto de reglas que defina nuestra gramática puede que no sea una base representativa, y por lo tanto sea muy probable que no tenga garantía de generar al lenguaje en su totalidad, sino sólo una parte. En cuanto, a esto Chomsky piensa que la razón principal por la que no se puede definir un conjunto finito de reglas para generar un lenguaje es que el lenguaje siempre está cambiando.
- Si ya existe una gramática para el lenguaje, Chomsky dice que es inadmisibles decidir si es un conjunto adecuado o no, pues no podemos probar cada una de las infinitas posibilidades.
- Evaluar las ventajas de las gramáticas existentes y crear una combinación de ellas para cubrir las necesidades en conjunto y así crear una mejor gramática.

Sin embargo, aun tomando en cuenta estas consideraciones, Chomsky (2005) se mantiene firme en que hay inconvenientes. Uno de ellos es que para cada individuo hay una gramática asociada, un subconjunto de la gramática del lenguaje que utiliza. Además, supone que aun y cuando el conjunto de reglas sólo pueda crear oraciones correctas, podría resultar incompleto. El lenguaje poético o literario no necesariamente sigue reglas, favorece la forma novedosa o diferente para la

comunicación de ideas. Al recurrir al conocimiento y expresividad del emisor y receptor, la interpretación es un nuevo factor a considerar, ya que se depende de la expresividad del emisor como en la capacidad de comprensión del receptor.

#### **2.2.4 Fonotáctica**

Una de las cualidades más llamativas en un lenguaje es la forma en que se habla un idioma.

La fonotáctica es una rama de la fonología que se ocupa de identificar las reglas de combinaciones en los sonidos de las palabras y la frecuencia de las sucesiones fonéticas en las lenguas, de acuerdo a Llorach (1994).

Continuamente se crean palabras nuevas o modifican las ya existentes, pero es curioso que, así como los seres vivos, sólo las más aptas sobrevivan. Esta supervivencia depende de la frecuencia con que se usen. Los sonidos que se forman al unir reglas y restricciones con fonemas en los idiomas es que hacen único al lenguaje, al darle una prosodia y musicalidad específica. Hay palabras que son naturales en un idioma, pero sencillamente imposibles o muy improbables en otro, por la diferencia entre la estructura de las palabras. Por ejemplo, cualquier palabra proveniente de otro idioma en Japón, de acuerdo a sus reglas de pronunciación, es muy probable que no sea pronunciable y sea necesario buscar una alternativa a cada uno de los sonidos.

De ahí la importancia de definir ciertas reglas en la definición de las palabras y no ser tan constrictivos en la definición de las reglas del nuevo lenguaje artificial, de acuerdo a Rosenfeld (2010).

Para medir el peso de una palabra se divide en partes: el núcleo, formado por una vocal, diptongo o triptongo; el ataque, formado a partir de diptongos y triptongos; y la coda, formado por una consonante.

Por ejemplo, en japonés no existen sonidos en los que dos o más consonantes estén juntas. En inglés no están permitidos los sonidos /kn/ y /qn/ al inicio de una palabra. En español siempre se escribe 'm' antes de 'p' o 'b'.

Con esta división y un peso silábico, para la mayoría de los lenguajes llamado mora, los lingüistas dividen las palabras en tres subclases:

- ligeras (monomoraicas)
- pesadas (bimoraicas)
- ultrapesadas (trimoraicas)

Estas divisiones surgen a partir de las siguientes reglas: Si hay una consonante al inicio de la sílaba, ésta no tiene peso. Tendrá un peso de uno si es corta. Tendrá un peso de dos si es larga.

En cuanto a la coda silábica varía según el lenguaje. Por ejemplo en Japón pesan uno, y en el inglés pesan dependiendo de la sílaba acentuada.

Como quedó anotado anteriormente, el lenguaje, al igual que los organismos vivos, está sometido a las fuerzas de la selección natural. Una de las fuerzas principales que selecciona los vocablos que se incorporarán a una lengua y se mantendrán a lo largo del tiempo es la facilidad con la que se pronuncia.

Las palabras que contravienen las reglas de la fonotáctica son más bien raras porque la selección sonora va en su contra.

El principio general del Alfabeto Fonético Internacional es otorgar a cada sonido un símbolo. Esto significa que el Alfabeto Fonético Internacional no usa combinaciones de letras a menos que el sonido representado pueda ser visto como una secuencia de dos o más sonidos. Ninguna letra tiene valores que dependan del contexto, como la *c* en la mayoría de los idiomas europeos.

**EL ALFABETO FONETICO INTERNACIONAL (actualizado en 2005)**

**CONSONANTES (INFRAGLOTALES)**

	LABIAL		CORONAL					DORSAL			RADICAL		GLOTALES
	BILABIAL	LABIODENTAL	DENTAL	ALVEOLAR	POSTALVEOLAR	RETROFLEJA	PALATAL	VELAR	UVULAR	FARINGEA	EPIGLOTALES		
NASAL	m	ɱ	n			ɳ	ɲ	ŋ	ɴ				
OCLUSIVA	p b	ɸ β	t d			ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ	ʕ	
FRICATIVA	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	ħ ʕ	h ɦ	
AFROXIMANTE		ʋ	ɹ			ɻ	j	ɰ					
VIBRANTE MÚLTIPLE	ʙ		r						ʀ		ʁ		
VIBRANTE SIMPLE		ʋ	ɹ			ɻ							
FRICATIVA LATERAL			ɬ ɮ			ʈ	ç	ʃ					
AFROXIMANTE LATERAL			l			ɭ	ʎ	ʟ					
VIBR. SIMPLE LATERAL			l			ɭ							

Las consonantes alineadas a la izquierda son sordas, las alineadas a la derecha sonoras. Las casillas en gris son articulaciones consideradas imposibles.

**CONSONANTES (SUPRAGLOTALES)**

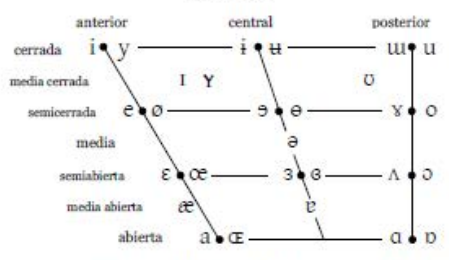
CLIC	IMPLICATIVA	EYECTIVA
◌̥ bilabial	ɸ bilabial	ʔ como en:
dental	ɖ dental / alveolar	pʰ bilabial
! (post)alveolar	f palatal	tʰ dental / alveolar
‡ palatoalveolar	ɣ velar	kʰ velar
lateral alveolar	ʁ uvular	sʰ fricativa alveolar

**CONSONANTES (COARTICULADAS)**

- ʌ fricativa labiovelar sorda
- ʋ aproximante labiovelar sonora
- ɰ aproximante labiopalatal sonora
- ç fricativa alveopalatal sorda
- ʝ fricativa alveopalatal sonora
- ʃ y x simultáneas

kp̥ tʃ̥ Africadas y dobles articulaciones pueden representarse con dos símbolos atados con una cuña

**VOCALES**



Las vocales a la izquierda del punto no son labializadas, las de la derecha son labializadas

**SUPRASEGMENTALES**

- ˈ acento principal
- ˈˈ acento extra
- ˌ acento secundario
- eː larga
- eˑ semilarga
- eˑ corta
- ˌ rotura silábica
- ˌ enlase
- ˌ ENTONACIÓN
- ˌ grupo menor (pie)
- ˌ grupo mayor (entonación)
- ˌ ascenso global
- ˌ descenso global

**TONO**

- NIVEL
- ˈ extra alto
- ˌ alto
- ˌ medio
- ˌ bajo
- ˌ extra bajo
- ˌ NIVELACIÓN
- ˌ -1 tono
- ˌ +1 tono
- ˌ ascendente
- ˌ descendente
- ˌ descendente alto
- ˌ descendente bajo
- ˌ ascendente
- ˌ descendente
- ˌ ascendente

**DIACRÍTICOS** En algunos pueden aparecer arriba: ̂ (tendencia fricativa), ̃ (sonora mate), ̄ (ataque glotal), ̅ (schwa epentético), ̆ (diphtongación)

SILABICIDAD Y TENDENCIA	FONACIÓN	ARTICULACIÓN PRIMARIA	ARTICULACIÓN SECUNDARIA
ɹ ɻ silábica	ɹ̥ ɹ̃	ɹ̥ ɹ̃ dental	ɹ̥ˠ ɹ̃ˠ labializada
ɹ̥ ɹ̃ no silábica	ɹ̥̃ ɹ̃̃	ɹ̥̃ ɹ̃̃ apical	ɹ̥̃ˠ ɹ̃̃ˠ palatalizada
ɹ̥ˠ ɹ̃ˠ aspirada	ɹ̥̃ˠ ɹ̃̃ˠ	ɹ̥̃ˠ ɹ̃̃ˠ laminal	ɹ̥̃ˠˠ ɹ̃̃ˠˠ velarizada
ɹ̥ˠˠ tendencia nasal	ɹ̥̃ˠˠ ɹ̃̃ˠˠ	ɹ̥̃ˠˠ ɹ̃̃ˠˠ avanzada	ɹ̥̃ˠˠˠ ɹ̃̃ˠˠˠ faringizada
ɹ̥ˠˠˠ tendencia lateral	ɹ̥̃ˠˠˠ ɹ̃̃ˠˠˠ	ɹ̥̃ˠˠˠ ɹ̃̃ˠˠˠ retraída	ɹ̥̃ˠˠˠˠ ɹ̃̃ˠˠˠˠ velarizada o faringizada
ɹ̥ˠˠˠˠ tendencia no audible	ɹ̥̃ˠˠˠˠ ɹ̃̃ˠˠˠˠ	ɹ̥̃ˠˠˠˠ ɹ̃̃ˠˠˠˠ centralizada	ɹ̥̃ˠˠˠˠˠ ɹ̃̃ˠˠˠˠˠ medio centralizada
ɹ̥̃̃ ɹ̃̃̃ descenso lingual (ɹ̃̃ es aproximante bilabial)	ɹ̥̃̃̃ ɹ̃̃̃̃	ɹ̥̃̃̃̃ ascenso lingual (ɹ̃̃̃̃ es fricativa alveolar sonora no silbante)	

Figura 2.10 Alfabeto fonético internacional, traducido al español. Se desglosa el sistema de notación estándar utilizado por lingüistas para representar los sonidos del habla de cualquier idioma de manera precisa y unívoca

## THE INTERNATIONAL PHONETIC ALPHABET (revised to 2020)

CONSONANTS (PULMONIC)

© 2020 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
◌ Bilabial	ɓ Bilabial	ʼ Examples:
Dental	ɗ Dental/alveolar	ɓʼ Bilabial
! (Post)alveolar	ɟ Palatal	tʼ Dental/alveolar
‡ Palatoalveolar	ɠ Velar	kʼ Velar
Alveolar lateral	ɠ Uvular	sʼ Alveolar fricative

OTHER SYMBOLS

ʌ Voiceless labial-velar fricative	ɕ ʑ Alveolo-palatal fricatives
ʋ Voiced labial-velar approximant	ɺ Voiced alveolar lateral flap
ɥ Voiced labial-palatal approximant	ɥ Simultaneous ʃ and x
ħ Voiceless epiglottal fricative	Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.
ʕ Voiced epiglottal fricative	
ʔ Epiglottal plosive	

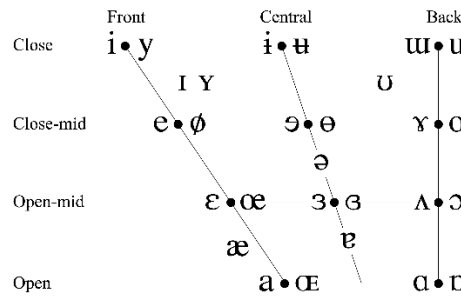
DIACRITICS

◌ Voiceless	◌̥ ◌̜	◌ Breathily voiced	◌̬ ◌̭	◌ Dental	◌̪ ◌̫
◌ Voiced	◌̦ ◌̧	◌ Creaky voiced	◌̨ ◌̩	◌ Apical	◌̽ ◌̾
◌ Aspirated	◌ʰ ◌̚	◌ Linguolabial	◌̼ ◌̻	◌ Laminal	◌̻ ◌̼
◌ More rounded	◌̙	◌ Labialized	◌̜ ◌̝	◌ Nasalized	◌̃
◌ Less rounded	◌̘	◌ Palatalized	◌̟ ◌̠	◌ Nasal release	◌̚
◌ Advanced	◌̟	◌ Velarized	◌̙ ◌̘	◌ Lateral release	◌̚
◌ Retracted	◌̠	◌ Pharyngealized	◌̙ ◌̘	◌ No audible release	◌̚
◌ Centralized	◌̠	◌ Velarized or pharyngealized	◌̙		
◌ Mid-centralized	◌̠	◌ Raised	◌̥ (ɹ̥ = voiced alveolar fricative)		
◌ Syllabic	◌̩	◌ Lowered	◌̜ (β̜ = voiced bilabial approximant)		
◌ Non-syllabic	◌̥	◌ Advanced Tongue Root	◌̟		
◌ Rhoticity	◌̠ ◌̡	◌ Retracted Tongue Root	◌̠		

Some diacritics may be placed above a symbol with a descender, e.g. ɲ̠̚

Typefaces: Doulos SIL (metatext); Doulos SIL, IPA Kiel, IPA LS Uni (symbols)

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

SUPRASEGMENTALS

ˈ Primary stress	ˈfounəˈtʃən
ˌ Secondary stress	
ː Long	eː
ˑ Half-long	eˑ
◌ Extra-short	ɛ̥
Minor (foot) group	
Major (intonation) group	
· Syllable break	·i.ækt
◌ Linking (absence of a break)	

TONES AND WORD ACCENTS

LEVEL	CONTOUR
ě or ǝ	↗ Extra high
é	↘ High
ē	↔ Mid
è	↘ Low
ě	↗ Extra low
↓	Downstep
↑	Upstep
↗	↗ Rising
↘	↘ Falling
↗↘	↗↘ High rising
↘↗	↘↗ Low rising
↗↘↗	↗↘↗ Rising-falling
↗↘↗↘	↗↘↗↘ Global rise
↘↘↗↘	↘↘↗↘ Global fall

Figura 2.11 Alfabeto fonético internacional actualizado. Se desglosa el sistema de notación estándar utilizado por lingüistas para representar los sonidos del habla de cualquier idioma de manera precisa y unívoca

## **2.2.5 Léxico**

El léxico de las lenguas se forma poco a poco por los diferentes tipos de contacto que tiene con su contexto, es decir, un lenguaje en la costa va a tener un diferente léxico a un lenguaje de las montañas. Las lenguas evolucionan y adquieren vocabulario a partir de cómo se vive.

### **2.2.5.1 Afiliación genética**

Son aquellas palabras que han estado presentes desde etapas muy tempranas. Estas palabras han evolucionado tanto fonéticamente que su origen, su raíz, no es siempre clara. En medio de ambas, hay toda una evolución y divergencias que surgen a partir de eventos que influenciaron a los hablantes de la lengua.

### **2.2.5.2 Identidad e integración**

Son palabras que han sido introducidas en la lengua, generalmente por escritores para la designación de conceptos novedosos o la búsqueda de algún fin estilístico. Estas palabras creadas no sufren las evoluciones fonéticas que suelen existir en la evolución natural de un idioma.

La adopción masiva de este tipo de léxico no es constante, sino que se ha concentra en épocas específicas, coincidiendo a menudo con eventos históricos de gran magnitud, tales como la revolución industrial o la ilustración.

### **2.2.5.3 Patrimonio y cultura**

Para algunas palabras patrimoniales, el desarrollo lingüístico se detiene en una fase muy temprana, casi al momento de su aparición. Algunas de las palabras dependen de la presión cultural, por ejemplo, en Japón se hace mucho hincapié en el respeto con el que se expresan las cosas. Otro ejemplo sería el de Polonia, en donde para ser respetuoso el emisor carga responsabilidad hacia él en sus acciones.

#### **2.2.5.4 Préstamos**

A cualquier lengua se le han ido incorporando préstamos léxicos de lenguas cercanas. Unos se fueron incorporando a partir del comercio, otras desde la guerra o de los avances tecnológicos, la necesidad para describir acciones u objetos. Por ello, son difíciles de reconocer. Anteriormente eran fundamentalmente de voces prerromanas, germanismos y arabismos; hoy en día mayormente de voces anglosajonas.

#### **2.2.5.5 Dialectos**

Los dialectos son pequeñas mutaciones que existen en una lengua a partir del lugar originario en donde se habla, es decir, las posibles variedades de una lengua; frecuentemente se usa el término dialecto para referirse a una variante geográfica de una lengua asociada con una determinada zona. Los cambios varían en significancia, a veces es sólo la diferencia entre pronunciaciones, en otros la diferencia radica en si existen o no la palabra para describir un objeto o actividades que se realicen con frecuencia. Por ejemplo, en una aldea cerca del río, se tendrán palabras para describir la pesca, los anzuelos, la carnada, las trampas; mientras que en las montañas se tendrán palabras para describir la minería, las herramientas, la oscuridad.

También son afectados por todo el contexto que los rodea. Un ejemplo general sería el del español, la evolución y el cambio que se ha generado a través de América sobre el español nos hace diferenciar entre cómo habla una persona en Tijuana a una en la Ciudad de México y otra en Yucatán.

### 2.2.5.6 Cambios de significado

Hay cuatro razones principales por las cuales las palabras cambian de significado:

- **Ocasionales.** Son significados que aparecen de forma espontánea en las palabras. En anécdotas de 1950, durante la posguerra, la palabra *haiga* hacía referencia a un coche grande. Los usuarios acostumbraban decir: "El coche más grande que haiga".
- **Históricas.** A partir de la historia es que se van agregando significado a las palabras. Por ejemplo, 'pluma' comenzó como referencia a la pluma de un ave y más tarde significó también bolígrafo.
- **Lingüísticas.** El contexto es un factor que puede alterar puntualmente la acepción de un término.
  - **Ampliación.** Las palabras pueden ser utilizadas en un contexto que normalmente no son utilizados.
  - **Especialización.** Se obtiene otro significado de acuerdo a contextos específicos.
- **Psicológicas.** Las palabras obtienen otro significado a partir de las creencias del lugar en donde se usa.

A su vez, existen formas de expresión que le atribuyen cierto significado a las palabras, sin embargo, requieren que exista cierta intención del escritor o hablante para situarlas en un contexto específico y puedan adquirir un nuevo significado:

- **La metáfora.** Se establece un vínculo semántico entre dos ideas a partir de una semejanza que puede ser tanto real como imaginaria.
- **La metonimia.** Entre las dos palabras hay una semejanza entre conceptos, pero estos no son usados de forma literal.

Por último, está la onomástica, un proceso especial en el que nombres propios pasan a ser nombres comunes.

## 2.2.6 Morfosintaxis

La morfosintaxis estudia las reglas y elementos que conforman una oración. Por una parte, la morfología estudia la tipología de las palabras, si estas son sustantivos, verbos, preposiciones, etc. Por otro lado, Llorach (1994) plantea que la sintaxis organiza cada uno de los elementos de la oración para que ésta tenga coherencia.

A nivel morfológico cada palabra está conformada por morfemas y lexemas y a nivel fonológico, de fonemas. A la unión de morfemas y lexemas se le llama monema.

Los morfemas se clasifican en:

1. Dependientes, aquellos que requieren aparecer unidos a un lexema.
  - **Flexivos**, cuya función es aportar información gramatical. Indican la concordancia de género y número en elementos nominales y especifican las categorías de modo, tiempo, número y persona en la flexión verbal.
  - **Derivativos**, que sirven para formar nuevas palabras o derivadas. Ocurre al unir a los lexemas prefijos o sufijos.
    - Prefijos, se agregan antes del lexema.
    - Sufijos, se agregan después del lexema.
2. Independientes, que no requieren estar unidos al lexema como en el caso de los artículos que modifican al sustantivo.

## 2.2.7 Semántica

En un inicio, como las comunidades humanas eran más reducidas bastaban unas cuantas palabras para comunicarse. Conforme las relaciones sociales se hicieron más complejas, los individuos empezaron a requerir palabras más complejas que simbolizaran ideas más elaboradas.

La semántica se encarga del significado, la interpretación de las palabras. Esto puede ser a partir de dos formas diferentes, a partir de la estructura de la palabra o a partir del contexto en el que se utilizan de acuerdo a las explicaciones de Llorach (1994).

## 2.3 Lenguaje artificial

Un lenguaje artificial, también llamado *conlang*, de acuerdo a Emrys (2006), es un lenguaje construido íntegramente por un ser humano, configurando de forma específica la fonología, gramática y vocabulario, sin que intervenga la evolución natural.

Cualquier persona puede crear un lenguaje desde cero, crear su propio código secreto, agregar trasfondo a historias, darles significado a nombres, lugares o incluso poemas de un mundo imaginario; para ser utilizado en programación, como ayuda para el aprendizaje de lenguaje máquina, para facilitar la comunicación entre comunidades o incluso para crear un lenguaje universal.

### 2.3.1 Lenguajes *a priori* y *a posteriori*

Un lenguaje *a priori* no se basa en ninguna lengua existente. En esta rama de lenguajes artificiales se encuentran generalmente los lenguajes creados como un experimento; como un proyecto artístico, que no está pensado para un uso cotidiano; como un lenguaje auxiliar, que pueda ayudar a comunidades con lenguajes con características en común; o como lenguajes comunitarios, para unificar lenguajes de una región.

Entre ellos se encuentran: unilingua, klingon, na'vi, valyrian, lojban, mannish, atlantean y láadan, entre muchos otros.

Un lenguaje *a posteriori* basa su gramática y palabras en otro lenguaje natural, ya sea que se tomen elementos prestados, se cree léxico sobre el actual lenguaje o se recorten características para controlar el crecimiento de un lenguaje.

Unos ejemplos son: Liberanto, Interlingua, Esperanto, Eurolengo, Globish y Folkspraak, lenguaje para sordos, entre otros.

### **2.3.2 Auxlangs**

Un *auxlang*, o lenguaje auxiliar internacional, es creado con la intención de facilitar la comunicación entre dos comunidades que hablen diferentes idiomas.

#### **2.3.2.1 Esperanto**

El esperanto fue desarrollado alrededor de 1887 por L.L. Zamenhof a partir de lenguajes romance, germánicas del oeste y eslavas. Él buscaba un uso más razonable de los afijos que permitiera eliminar el número de raíces necesarias para comunicarse en su natal Polonia. Vale la pena mencionar que su uso se intensificó en tiempos de guerra, cuando individuos de diferentes nacionalidades en conflicto querían interactuar entre sí. Actualmente es hablado alrededor del mundo, con un estimado de cien mil a dos millones de personas.

Se suele resumir su gramática básica en dieciséis reglas, aquí están las diez más importantes:

1. El artículo indefinido no existe; sólo existe el artículo 'la', igual para todos los géneros, números y casos.

2. El nombre sustantivo termina en ‘-o’. Para formar el plural se añade la terminación ‘j’. La lengua sólo tiene dos casos: nominativo y acusativo. El último se forma añadiendo una ‘n’ al nominativo.
3. El adjetivo termina en ‘a’. Sus casos y números se forman como en los sustantivos. El grado comparativo se forma con la palabra ‘pli’ y el superlativo con la palabra ‘plej’.
4. Los pronombres personales son: mi, vi, li, ŝi, ĝi, si, ni, ili, oni.
5. El verbo es invariable en las personas y en los números. El presente termina en ‘as’, el pasado en ‘is’, el futuro en ‘os’, el condicional en ‘us’, el imperativo en ‘u’ y el infinitivo en i.
6. El adverbio termina en ‘e’. Sus grados de comparación se forman como los del adjetivo.
7. Todas las preposiciones rigen, por sí mismas, el nominativo.
8. Toda palabra se pronuncia del mismo modo como se escribe.
9. Las palabras compuestas se forman por la simple reunión de los elementos que las forman. En ellas la palabra fundamental va siempre al final.
10. Toda preposición tiene, en esperanto, un sentido invariable y bien determinado, que fija su empleo.

### **2.3.2.2 Interlingua**

Fue desarrollado entre 1937 y 1951 por la Asociación de la Lengua Auxiliar Internacional, utilizando una gramática simple y regular a partir de los idiomas inglés, francés, italiano, latín, portugués, y español; además del vocabulario que tienen los países en común, es lo que

la hizo fácil de aprender, por lo menos para aquellos que hayan sido originarios de alguna lengua en las que se basó. Actualmente hay entre mil y quince mil hablantes alrededor del mundo.

Las principales bases gramaticales son:

1. El artículo determinado es 'le' para todos los géneros y números. El artículo indeterminado para todos los géneros es 'un'.
2. Los sustantivos que terminan en vocal forman el plural por la adición de la terminación '-s', y los sustantivos que terminan en una consonante forman el plural por la adición de '-es'.
3. El adjetivo es invariable, es decir, no concuerda con el sustantivo ni con el pronombre. Puede estar delante o detrás del sustantivo.
4. El adverbio se forma con la adición de '-mente' al adjetivo. El comparativo de los adjetivos y adverbios se forma con 'plus' y el superlativo por 'le plus', y para grados de inferioridad por 'minus' y 'le minus'.
5. El verbo tiene solamente siete formas diferentes: -, -va, -ra, -rea, -r, -nte, -te.
6. Hay tres infinitivos: -ar, -er, ir.
7. La misma forma del verbo se usa para todas las personas singulares y plurales. Se omite el pronombre solamente en el imperativo y en ciertas expresiones impersonales.

### **2.3.2.3 Lengua de señas**

De acuerdo al Instituto De Las Personas Con Discapacidades (2017), el lenguaje de señas es un sistema lingüístico complejo que se articula mediante signos gestuales manuales, los

cuales se complementan con componentes no manuales como expresiones faciales, la mirada y el movimiento corporal. Poseen una riqueza léxica y complejidad gramatical equiparable a la de cualquier lengua oral.

El manual continua, definiéndolo como un lenguaje natural con una gramática que permite expresar el pensamiento a través gestos. Además de que se trata también de la lengua materna del sordo y su forma natural de comunicación. Al igual que cualquier otro idioma, varía de una comunidad a otra, el tener y evolucionar su propio léxico;

Su estructura es completa, tiene morfología, sintaxis, gramática, léxico y también rasgos idiosincráticos y de identidad a nivel regional, como lo sería el acento en las lenguas orales.

Sin embargo, actualmente, carece de remodificación escrita, es decir, una forma escrita estándar.

Rastrear la evolución de la Lengua de Señas Mexicana resulta una tarea complicada debido a la carencia de registros fiables y datos exactos, a causa de los años en que estuvo prohibida.

Se estima que actualmente hay al menos unos 138 tipos de lenguas de señas en el mundo.

Se estructura mediante el uso del cuerpo en el espacio. Dentro de este espacio proyectado existen tres límites: el vertical, el horizontal y un tercero que indica la proximidad de las manos con respecto al cuerpo. El límite vertical va de la cintura a la coronilla de la cabeza; el límite horizontal, hasta la altura de los codos, con los brazos doblados. Cualquier movimiento que sobrepase estas dimensiones será interpretado como exageración o énfasis.

Las señas se dividen en cuatro tipos diferentes. Esto depende del uso de una o dos manos o de su movimiento, sea simultáneo o no:

- Una seña manual se articula con una sola mano.

- Una seña bimanual se articula con dos manos a la vez de manera dispar, libre o asimétrica.
- Una seña simétrica se articula con dos manos a la vez, siempre a través de movimientos en espejo, idénticos e inversos (alternados) o inversamente proporcionales (simultáneos).
- Una seña compuesta se articula a través de al menos dos señas simples o tres configuraciones distintas.

**Seña:** SC: I. SM; II. SM

**Configuración:** I. B-P.1, II. 2.1

**Orientación:** I. Palma hacia adentro; II. Palma hacia la izquierda.

**Ubicación:** I. Sobre el pecho; II. A la altura de la cara del lado derecho.

**Movimiento:** I. y II. Recto hacia enfrente.

**Definición:** 1. Colonia de la Ciudad de México ubicada en la delegación Cuauhtémoc. 2. Estación del STC Metro de la línea B.



Figura 2.12 Entrada del diccionario de Lengua de Señas Mexicana, descripción de la seña, apoyo visual y análisis lingüístico.

**Seña:** SB

**Configuración:** MD L.1, MB S.1

**Orientación:** MD palma hacia la izquierda, MB palma hacia la derecha.

**Ubicación:** La MD inicia a la altura de la cabeza y termina sobre MB. MB a la altura del pecho.

**Movimiento:** La MD se mueve formando un arco y golpea la MB.

**Rasgos no manuales:** Cejas hacia arriba, boca abierta.

**Definición:** 1. v. tr. Hallar o encontrar algo o a alguien que no se conocía. 2. v. prnl. Darse cuenta clara de algo, tomar conciencia de ello.



Figura 2.13 Entrada del diccionario de Lengua de Señas Mexicana, descripción de la seña, apoyo visual y análisis lingüístico.

### 2.3.3 Engelang

Son lenguajes contruidos para experimentar o probar hipótesis de cómo funcionan o podrían funcionar los lenguajes, como tal no tienen un fin general de ser utilizados por la comunidad en general, pero sí el de intentar cumplir el objetivo para el que fue creado. Se dividen en tres subcategorías: lógicos, filosóficos y experimentales.

Los lenguajes lógicos intentan reducir la ambigüedad semántica en las oraciones al mínimo. Los más conocidos son: Loglan y Lojban. Éstas se basan en premisas de las que sólo se pueden obtener respuestas verdaderas o falsas, por lo que se consideran como una herramienta posible para comunicarse más fácilmente con programas y máquinas.

#### 2.3.3.1 *Loglan*

Desarrollado en 1955 por el doctor James Cooke Brown, el *Loglan* fue diseñado con el fin de investigar la hipótesis Sapir-Whorf, la cual propone que dependiendo del lenguaje que uno aprende es la forma en la que percibimos el mundo. Se esperaba que con este lenguaje tan lógico las personas empezaran, además de hablar, a pensar de forma lógica.

En el *loglan* se pueden encontrar reglas a partir de los tres tipos de palabras:

- Predicados o palabras de contenido. Son la mayoría de las palabras, éstas conllevan significado. Se posicionan antes de la expresión que lo modifica. También pueden ser usados de forma adverbial para modificar el predicado principal. Además de que pueden ser modificados para indicar el tiempo de la acción.

- Palabras de estructura o palabras pequeñas. Son palabras que modifican al predicado o muestran cómo se relacionan unas con otras.
- Nombres. Estos siempre terminan en consonante, al contrario de las estructuras y predicados, que terminan en vocal.

El principio de los lenguajes filosóficos es el trascender a una verdad mística a partir de los principios más básicos, es decir, las palabras se crean a partir de muy pocas raíces que consideran fundamentales. Sin embargo, aun y cuando la invención de estos lenguajes es intentar llegar a la perfección, las lenguas modernas como el toki pona entienden que ese tipo de aseveraciones son absurdas.

### **2.3.3.2 Toki pona**

Es un lenguaje minimalista diseñado por Sonja Lang e inspirado por la filosofía taoísta que utiliza muy pocas raíces para crear palabras. Se concentra en los conceptos universales de los lenguajes y apenas cuenta con catorce fonemas y ciento veinte raíces de palabras. Fue publicado por primera vez en el 2001, pero completado hasta el 2014.

En cuanto a las reglas figuran:

- Las oraciones estarán estructuradas en una sintaxis SVO.
- El sujeto y el predicado son separados por la palabra 'li'.
- La letra 'e' siempre va antes de un objeto directo.
- El objeto directo siempre va antes de las frases preposicionales en un predicado.
- Los verbos complejos están separados de la oración principal por un 'la'.

El vocabulario de ciento veinte palabras está diseñado a partir del concepto de vida simple, sin ninguna complicación de la vida moderna, pero con este número de palabras basta para describir la mayoría de las ideas que se deseen expresar.

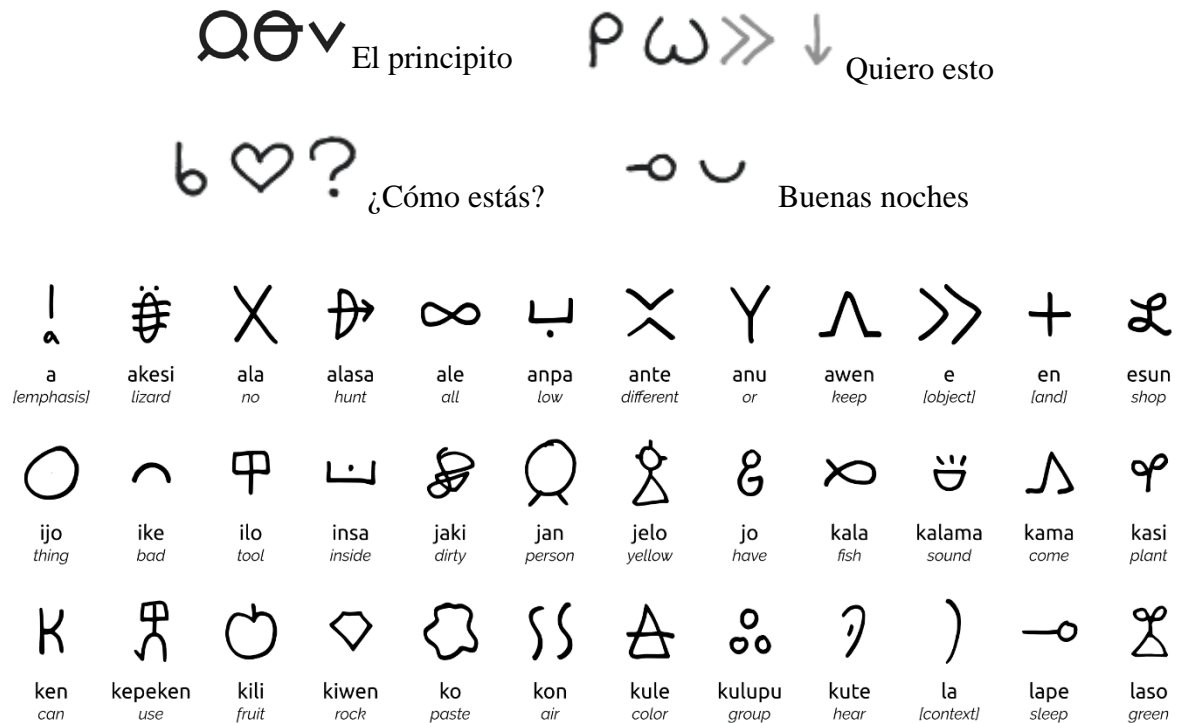


Figura 2.14 Treinta y seis símbolos de los ciento veinte símbolos que componen el lenguaje Toki Pona

En los lenguajes experimentales el objetivo es el de investigación. En su mayoría estos lenguajes se crean para experimentar nuevas formas de comunicación. La mayoría de estos lenguajes cuestiona qué es lo que cada palabra nos hace pensar, cómo es que nuestro cerebro procesa la información de cierta manera.

### 2.3.4 artlangs

Los lenguajes artísticos son construcciones diseñadas para el entretenimiento, películas, videojuegos y libros. Estos lenguajes tienden a ser irregulares en sus gramáticas, haciéndolos

parecer más naturales con respecto de otro tipo de lenguajes artificiales. Sin embargo, estos lenguajes completos sólo son una pequeña parte, muchos de estos lenguajes no son completamente funcionales y sólo sirven para un cierto número de casos especiales. El objetivo de estos lenguajes casi siempre es el de fortalecer algún mundo ficticio, como lo hizo J. R. R. Tolkien.

### 2.3.4.1 Lenguajes élficos

J. R. R. Tolkien empezó a construir su primera lengua élfica en 1910. A partir de su gusto por los lenguajes y su motivación por crear una mitología, fue que decidió crear una historia para estas lenguas élficas que tenía en mente. Tal como en los procesos naturales, Tolkien hizo revisiones en su gramática y vocabulario constantemente. El cambio constante en la historia de los pueblos élficos influenciaba la evolución del lenguaje de sus habitantes.

Para la construcción de sus lenguajes, Tolkien usaba un modelo de árbol y una filología comparativa. Comenzaba con un sistema fonético del proto-lenguaje y continuaba con los lenguajes derivados con los cambios necesarios.

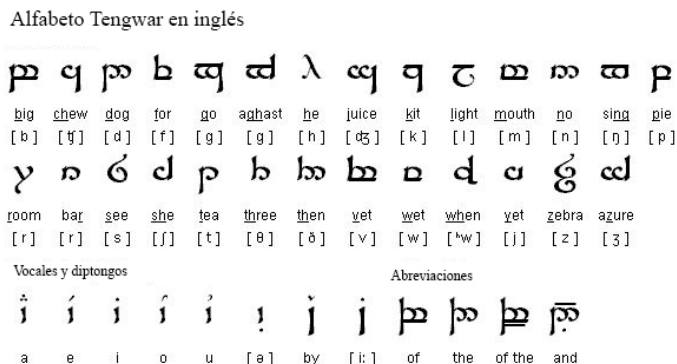


Figura 2.15 Guía para escribir en inglés utilizando el sistema de escritura Tengwar, creado por J.R.R. Tolkien

En una entrevista J. R. R. Tolkien dijo:

La invención de las lenguas es la fundación. Las “historias” se hicieron más bien para proporcionar un mundo para las lenguas y no al contrario. Para mí un nombre viene primero y la historia sigue. Hubiera preferido escribir en Élfico. Pero, claro, una obra como El Señor de los Anillos ha sido editada y sólo se ha dejado tanto “lenguaje” como pensé que sería digerido por los lectores. (Ahora me parece que a muchos les hubiera gustado más.) Pero hay una gran cantidad de materia lingüística (aparte de los nombres y palabras “élficas”) incluidos o mitológicamente expresados en el libro. De todos modos, para mí es en gran parte un ensayo de "estética lingüística", como a veces digo a las personas que me preguntan “¿de qué se trata?”<sup>1</sup>

El problema de especificar una gramática para los lenguajes élficos es que J. R. R. Tolkien nunca publicó en vida artículos que hablaran de ella. Sin embargo, a partir de publicaciones hechas por *Vinyar Tengwar* es que se ha comprendido un poco más de cómo funcionaban, además de contabilizar más de veinticinco mil palabras.

---

<sup>1</sup> *The invention of languages is the foundation. The “stories” were made rather to provide a world for the languages than the reverse. To me a name comes first and the story follows. I should have preferred to write in “Elvish”. But, of course, such a work as The Lord of the Rings has been edited and only as much “language” has been left in as I thought would be stomachable by readers. (I now find that many would have liked more.) But there is a great deal of linguistic matter (other than actually “elvish” names and words) included or mythologically expressed in the book. It is to me, anyway, largely an essay in “linguistic aesthetic”, as I sometimes say to people who ask me “what is it all about?”*

### 2.3.4.2 *klíngon*

Marc Okrand diseñó el *klíngon* para la serie de televisión *Star Trek* después de que se escuchara por primera vez en la película *Star Trek* en 1979. Este lenguaje es bastante específico, su vocabulario se centra mucho en cómo describir el espacio y la guerra intergaláctica, por lo que su uso para describir tareas cotidianas es un tanto difícil.

### 2.3.4.3 *trunic*

En el juego *Trunic*, desarrollado por el Isometricorp Games se creó en parte un *conglang* (solamente se crearon los símbolos y se les asignaron sonidos del inglés) para profundizar en el universo de fantasía en ruinas.

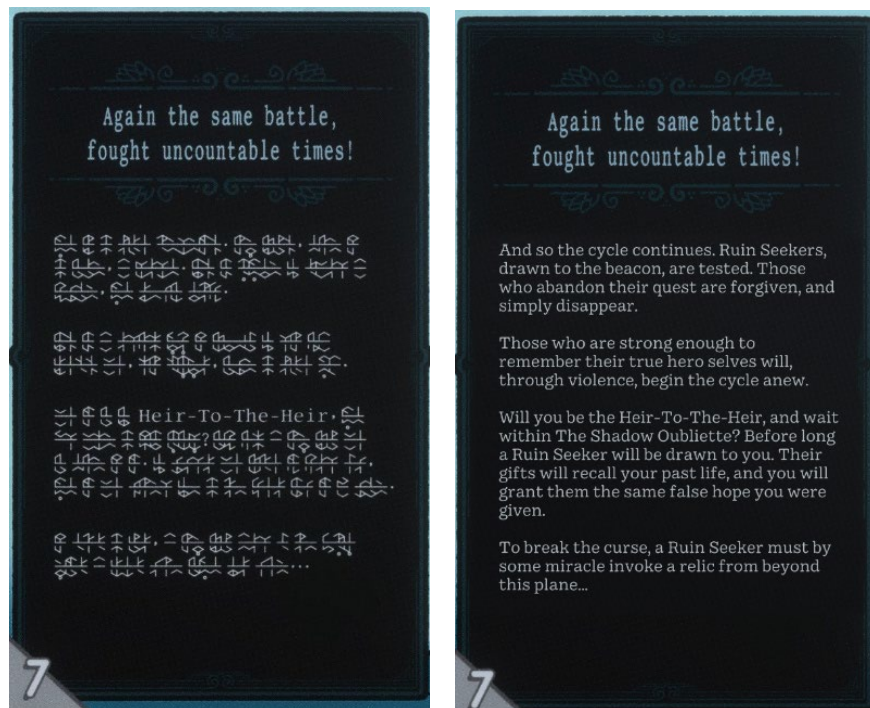


Figura 2.16 Leyenda de Trunic, comparativa entre Trunic e inglés.

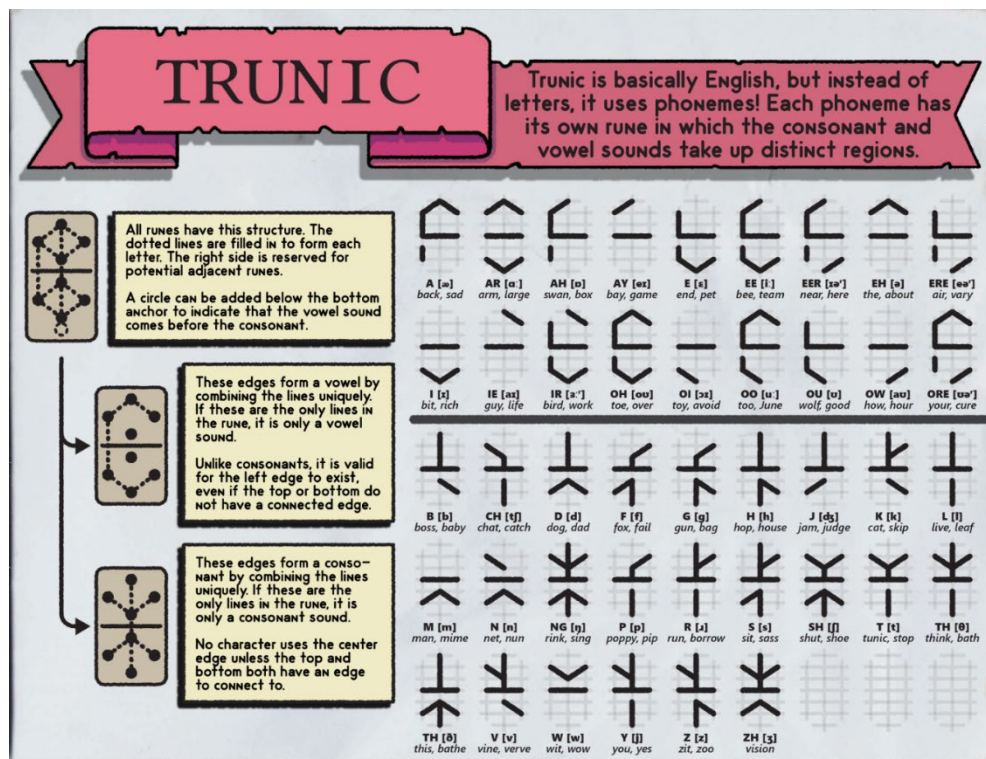


Figura 2.17 Guía para escribir en inglés utilizando el sistema de escritura Trunic

### 2.3.5 Lista de palabras

La lista *swadesh* fue desarrollada por el lingüista Morris Swadesh entre 1940 y 1950. En ella se incluyen partes del cuerpo, verbos, fenómenos naturales, etc. Términos universales que todo lenguaje contiene con los que se podría tener la más básica comunicación sin conocimientos de sintaxis o gramática. Por ello, los lingüistas la utilizan a la hora de crear un lenguaje artificial.

1. I	34. horn	67. lie
2. you	35. tail	68. sit
3. we	36. feather	69. stand
4. this	37. hair	70. give
5. that	38. head	71. say
6. who?	39. ear	72. sun
7. what?	40. eye	73. moon
8. not	41. nose	74. star
9. all	42. mouth	75. water
10. many	43. tooth	76. rain
11. one	44. tongue	77. stone
12. two	45. claw	78. sand
13. big	46. foot	79. earth
14. long	47. knee	80. cloud
15. small	48. hand	81. smoke
16. woman	49. belly	82. fire
17. man	50. neck	83. ash

18. person	51. breast	84. burn
19. fish	52. heart	85. path
20. bird	53. liver	86. mountain
21. dog	54. drink	87. red
22. louse	55. eat	88. green
23. tree	56. bite	89. yellow
24. seed	57. see	90. white
25. leaf	58. hear	91. black
26. root	59. know	92. night
27. bark	60. sleep	93. hot
28. skin	61. die	94. cold
29. flesh	62. kill	95. full
30. blood	63. swim	96. new
31. bone	64. fly	97. good
32. grease	65. walk	98. round
33. egg	66. come	99. dry
		100. name

*Tabla 2.1 Lista de palabra swadesh de cien palabras.*

# 3 Desarrollo

## 3.1 Diseño del algoritmo

### 3.1.1 Planteamiento del problema

La búsqueda por ampliar las formas de comunicación es inherente al ser humano. Ya sea a través de la creación de gestos, música, imágenes, palabras e incluso en el desarrollo de lenguajes enteros, para simplificar la comunicación o para explorar modos de expresión que evoquen emociones distintas. Por ejemplo, el desarrollo de la notación matemática fue una extensión de estas formas de comunicación. Sin embargo, aunque nuestro conocimiento sobre el lenguaje es bastante extenso, todavía entre lingüistas se habla sobre teorías del lenguaje. A lo largo de los años se ha buscado un lenguaje definitivo, un lenguaje que trascienda las barreras de nuestros lenguajes actuales. Los intentos por expandir las formas de comunicación y entenderlas no se detienen.

En los últimos años los lingüistas han intentado atacar los problemas de comunicación entre individuos de diferentes países, profundizar en las formas de comunicación del ser humano con su entorno, avanzar en el campo tecnológico e incluso crear un lenguaje para ser utilizado en entretenimiento. Por tal motivo, los lingüistas han creado lenguajes como el esperanto, el lojban y el toki pona, entre muchos otros, para enriquecer y ampliar las diferentes áreas de aplicación. No obstante, el tiempo y dedicación que requiere el desarrollo de un solo lenguaje es sumamente demandante.

Este trabajo propone acelerar el proceso de desarrollo en el que se crea un lenguaje artificial mediante la aplicación del cómputo evolutivo.

### 3.1.2 Hipótesis y soluciones esperadas

A partir del cómputo evolutivo se desarrollará de forma básica un nuevo lenguaje artificial que podrá ser usado como *artlang*. Éste quedará como base para su expansión. Como solución se espera un lenguaje simple desarrollado en dos partes: a partir del algoritmo genético, una lista *swadesh* con cien palabras; y una gramática definida a partir del algoritmo extendido para *LR parsers* definido en la tesis de Masaru Tomita (1984), “LR Parsers For Natural Languages”, tomando como base un *treebank* de cincuenta frases que quisiera que se pudieran expresar en este nuevo *artlang*. Este *treebank* contiene frases en español, inglés, japonés, latín y chino.

## 3.2 Primera fase de desarrollo

En la primera fase de desarrollo se programó el generador de palabras a partir del algoritmo genético simple, en el lenguaje de programación python y en el lenguaje javascript.

La definición del genotipo es en ambos lenguajes una clase con los siguientes elementos:

- Cromosoma, la palabra individuo en la que cada carácter es un gen.
- Punto de cruza, el que se utilizará al momento de cruza
- Aptitud, a partir de la evaluación de cada letra respecto a su posición
- Primer progenitor, primera palabra utilizada para la creación de este individuo
- Segundo progenitor, segunda palabra utilizada para la creación de este individuo



Figura 3.1 Componentes de un individuo: cromosoma, aptitud, punto de cruza, progenitor uno y progenitor dos

Las poblaciones estaban restringidas por valores definidos como los siguientes: generación máxima, seiscientos, el máximo de individuos por población, doscientos y un número de individuos élite, veinte. Esta porción élite fue definida porque no todos los individuos logran tener la aptitud mínima requerida para sobrevivir. Y el número máximo de generaciones se elige como freno en el momento aproximado en el que la población llega a un estado estable. Primeramente, se utilizó una población máxima de cien, pero el promedio de las aptitudes de los individuos nunca llegaba a ser alta, o menos de la mitad de los individuos llegaba a sobrevivir, por eso se extendió el número de individuos posibles: para que, si la mitad sobrevive, sea un número suficiente de palabras.

En cuanto al número de generaciones, se probó con una población de ochocientos y luego de mil doscientos, en estos casos el problema fue que se degradaba la aptitud general de las generaciones y en ciertos casos se perdían palabras interesantes, cuyos sonidos y estructura evocaban un significado concreto.

Los individuos están restringidos por longitud mínima, dos, y máxima, siete. Esto con la intención de que el programa no genere palabras demasiado largas o demasiado cortas, para que el lenguaje generado sea rápido de hablar, pero que las palabras sean lo suficientemente largas como para que no sean uno o dos sonidos por palabra. Dentro de este algoritmo se encuentra el algoritmo para generar los primeros individuos para la primera población.

---

**Algoritmo 2** Generación del cromosoma individuo

---

**Requiere:** Información de probabilidad de abecedario

- 1: Inicializa arreglo de individuo vacío
  - 2: Selección aleatoria de longitud de individuo
  - 3: **mientras** la cadena del individuo no mida la longitud seleccionada **hacer**
  - 4:     Girar ruleta para seleccionar una probabilidad
  - 5:     Agregar el carácter a la cadena que fue seleccionado
  - 6: **terminar mientras**
- 

Cada carácter es seleccionado cuando la probabilidad cae dentro del límite superior, si la probabilidad resultante del método es 0.0692, el carácter seleccionado sería B; si la probabilidad es de 0.868, el carácter seleccionado sería U.

Carácter por probabilidad				
A - 0	B - 0.05	C - 0.1	D - 0.15	E - 0.20
F - 0.25	G - 0.30	H - 0.35	I - 0.40	K - 0.45
L - 0.50	M - 0.55	N - 0.60	O - 0.65	Q - 0.70
R - 0.75	S - 0.80	U - 0.85	V - 0.90	X - 0.95
Z - 1				

*Tabla 3.1 Referencia de probabilidad acumulada para la selección de caracteres en el lenguaje*

A partir del alfabeto fonético se definieron los pesos para cada una de las letras del abecedario que conocemos. Para los pesos se consideraron estructuras fonotácticas a partir de la estructura de palabras en español, portugués y japonés, promoviendo aquellas combinaciones de sonidos que fueran sencillas y no chocantes o estridentes al momento de pronunciarlas. En ambos lenguajes de programación se definieron diccionarios con el peso de cada una de las letras, excluyendo las letras J, Ñ, P, T, W e Y.

Valores del abecedario				
A - 1	B - 1	C - 2	D - 2	E - 3
F - 2	G - 1	H - 3	I - 1	K - 2
L - 2	M - 1	N - 2	O - 3	Q - 1
R - 3	S - 3	U - 1	V - 3	X - 3
Z - 2				

Tabla 3.2 Referencia de valor de importancia para la evaluación

Además, se realizó una subvaloración de las letras para situaciones específicas de la fonotáctica.

Consonantes				
B - 3	C - 5	D - 3	F - 4	G - 1
H - 5	K - 2	L - 2	M - 1	N - 2
Q - 1	R - 4	S - 3	V - 5	X - 4
Z - 2				

Vocales				
A - 3	E - 4	I - 1	O - 5	U - 2

Tabla 3.3 Referencia de valor para casos especiales en los que un tipo de letra específico necesite ser seleccionada

Esto se hizo para ofrecer menor diversidad a la población de individuos en situaciones en donde las sílabas podrían caer en combinaciones no tan deseadas. En momentos en que la fonotáctica defina una palabra de cinco letras; sirva de ejemplo CVCVV, donde la segunda sílaba será evaluada con distintos valores para acentuar el uso de ciertas consonantes y vocales.

También se consideraron diferentes estructuras fonotácticas diseñadas a partir de la estructura de palabras en español, portugués y japonés, buscando que la combinación de sonidos entre las palabras no sea chocante o estridente al habla. La longitud de cada cadena es un factor que define el sonido de una palabra debido a la diferente longitud, estructura, pero en especial al peso mora en las sílabas de cada una de las posibles palabras. Es decir, una consonante está más presente y define en mayor medida el sonido de una palabra entre más corta ésta sea. Por ello, el número de consonantes definido en la fonotáctica de las palabras es mínimo: dan forma al sonido, pero no son parte de él.

Número de caracteres	Fonotáctica	Excepciones aceptadas
2	CV	
3	CVV	CCV
4	CVCV	VVCV
5	CVCVV	CVCVC y CVCVV
6	Se permitió cualquier tipo de fonotácticas, siempre y cuando no haya tres vocales continuas.	
7		

*Tabla 3.4 Reglas fonotácticas para la evaluación de palabras*

Para las palabras de dos caracteres se definió una fonotáctica de la forma CV. Con respecto de las palabras de tres caracteres, se permitieron fonotácticas de la forma CVV, aunque se permitieron excepciones de la forma CCV. En palabras de cuatro caracteres se permitieron fonotácticas de la forma CVCV, con la excepción en la forma VVCV. En las palabras de cinco caracteres se permitieron fonotácticas de la forma CVCVV, con excepciones de la

forma CVCVC y CVCVV. Y por último en las palabras de seis y siete caracteres, se permitió cualquier tipo de fonotácticas, siempre y cuando no haya tres vocales continuas.

En todos los casos en los que el individuo no tenga una longitud mencionada anteriormente, contenga tres consonantes o tres vocales juntas, se considera que el individuo muere. Su aptitud será cero, porque en este contexto no tiene las características para sobrevivir.

Antes de continuar, el programa tiene que evaluar la primera generación, obtener la aptitud de cada uno de los individuos. Para realizarlo utiliza la función general del programa.

---

**Algoritmo 3** Evaluación del individuo: para obtener su aptitud

---

**Requiere:** Los diccionarios de equivalencias entre carácter y aptitud

- 1: Inicializar la aptitud de la población
- 2: Identificar longitud de individuo
- 3: **switch** (Longitud del individuo) :

- 4: **caso** Longitud entre dos y siete:

La aptitud es evaluada por posición, comparada contra el diccionario y sumada al individuo

- 5: **caso** En otro caso:

La aptitud asignada será cero.

- 6: **terminar switch**
- 

A partir de este momento el programa comienza el ciclo generacional, definido por el siguiente algoritmo.

---

**Algoritmo 4** Evolución de población

---

**Requiere:** Información de población e individuos

- 1: Inicializa arreglo de población vacío
  - 2: **mientras** la población no tenga el número de individuos definido **hacer**
  - 3:     Girar ruleta para seleccionar individuo uno y dos
  - 4:     Realizar cruza y mutación entre individuos seleccionados
  - 5: **terminar mientras**
-

Para generar una segunda generación y subsecuentes, los individuos son elegidos de forma aleatoria para competir en un torneo y decidir cuáles son más aptos para generar descendencia.

En este proceso es donde se utilizan los operadores de cruce y mutación.

En total se seleccionan cuatro individuos para comenzar el proceso de cruce: dos torneos de aptitudes, dos ganadores. En caso de que en algún torneo dos de los individuos a competir tengan la misma aptitud, sólo uno de ellos será seleccionado de forma aleatoria.

---

**Algoritmo 5** Cruza y mutación de individuos

---

**Requiere:** Arreglo de nueva población e información de individuos

- 1: Obtener puntos de cruce
  - 2: **si** cruce ocurre **entonces**
  - 3:     Definir arreglo A y arreglo B e inicializarlos vacíos
  - 4:     En arreglo A copiar desde el primer carácter hasta el punto de cruce del progenitor A
  - 5:     En arreglo B copiar desde el primer carácter hasta el punto de cruce del progenitor B
  - 6:     En Arreglo A copiar desde el punto de cruce del progenitor B hasta el final de su cadena
  - 7:     En Arreglo B copiar desde el punto de cruce del progenitor A hasta el final de su cadena
  - 8:     **para** los cromosomas nuevos **hacer**
  - 9:         Mutar carácter por carácter
  - 10:     **terminar para**
  - 11:     Introducir individuos a nueva población
  - 12: **terminar si**
- 

Estos individuos entonces pasan por el proceso de cruce. Se seleccionan dos puntos de cruce, uno para cada individuo, entre el primer carácter y el último. Si ocurre que el punto de cruce es el primer carácter, todo el individuo será sobrescrito, y no hay error en ello, es parte de la selección de genes dominantes. Una vez que se obtienen las nuevas cadenas, éstas, carácter

por carácter, pasan por un proceso de mutación. En cada uno de sus caracteres se aplica una función de mutación de muy baja probabilidad. Aquí no hay restricción en el carácter resultante, puede ser cualquier letra del abecedario. Al final de este proceso se han generado dos nuevos individuos, dos descendientes, que son introducidos en la nueva generación.

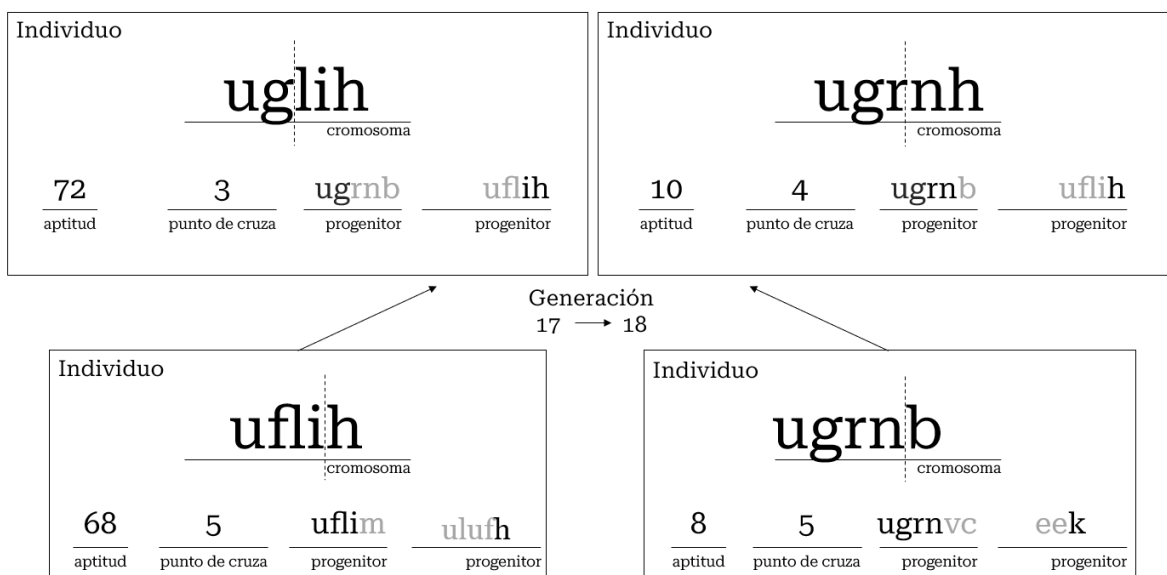


Figura 3.2 Apoyo visual de la reproducción de dos individuos, los dos nuevos individuos y de la posible cruce y mutación que ocurre a partir de los parámetros de cada uno de los individuos.

A lo largo de las generaciones aquí surge un problema: como no hay restricciones en la aptitud de los individuos, las funciones de evaluación no cambian y los individuos más aptos tienen menos dificultad en el duelo de selección para cruce, las palabras tienden a maximizar su aptitud. Después de un número pequeño de generaciones, los individuos terminan con el mismo cromosoma, son la misma palabra. Para corregir esto, se introdujo una función en el algoritmo para eliminar estos individuos duplicados e introducir nuevos individuos. Esto hizo dos cosas: mejoró los resultados, la variabilidad de individuos es mucho mayor, y redujo la velocidad con la que las palabras incrementan su aptitud.

---

**Algoritmo 6** Purgar individuos iguales y crear nuevos individuos

---

- 1: crear arreglo vacío llamado Visto
  - 2: **para** cada individuo en poblacion **hacer**
  - 3:     **si** individuo no existe en Visto **entonces**
  - 4:         Agregar a Visto
  - 5:     **terminar si**
  - 6: **terminar para**
  - 7: **mientras** Visto tenga menores individuos que el número de población **hacer**
  - 8:     Agregar individuos nuevos
  - 9: **terminar mientras**
  - 10: Evaluar nuevos individuos
- 

Ahora sí, una vez que los individuos de la generación han sido purgados, se evalúa cada uno de los individuos. Para ello se programaron siete funciones diferentes.

```
def thr(tmp):
    acc = 10
    if (tmp[0] in cons):
        acc = acc + cons[tmp[0]]*8
    else:
        acc = acc + abcval[tmp[0]]*3
    if (tmp[1] in voca):
        acc = acc + voca[tmp[1]]*9
    else:
        acc = acc + abcval[tmp[1]]*5
    if (tmp[2] in voca):
        acc = acc + voca[tmp[2]]*7
    else:
        acc = acc + abcval[tmp[2]]*3
    acc = ckt(tmp,acc)
    return acc

def go(tmp):
    acc = 50
    if (tmp[0] in cons):
        acc = acc + cons[tmp[0]]*9
    else:
        acc = acc + abcval[tmp[0]]*2
    if (tmp[1] in voca):
        acc = acc + voca[tmp[1]]*9
    else:
        acc = acc + abcval[tmp[1]]*2
    if (tmp[2] in cons):
        acc = acc + cons[tmp[2]]*8
    else:
        acc = acc + abcval[tmp[2]]*2
    if (tmp[3] in voca):
        acc = acc + voca[tmp[3]]*6
    else:
        acc = acc + abcval[tmp[3]]*4
    if (tmp[4] in voca):
        acc = acc + voca[tmp[4]]*5
    else:
        acc = acc + abcval[tmp[4]]*5
    acc = ckt(tmp,acc)
    return acc
```

Figura 3.3 Ejemplo de las funciones de evaluación dentro del programa para las diferentes longitudes, en este caso, longitud tres y longitud cinco.

Se tuvieron consideraciones para las palabras de longitud menor. Al momento de ser evaluadas son éstas las que, por simple definición, no pueden alcanzar aptitudes altas. Para

ello se les asigna una aptitud inicial. Esta aptitud inicial no busca que sean las mejores palabras, que puedan competir con palabras de siete caracteres, pero sí que por lo menos tengan un valor mínimo alto y puedan sobrevivir algunos de los duelos de selección para cruza.

Una vez evaluada la población, el programa comienza de nuevo a partir del Algoritmo 4, seleccionando individuos, realizando la cruza de ellos y, en algunos casos, mutación.

### Ejemplo de la evaluación

Un resultado preliminar de la lista de palabras ofrece los siguientes valores:

Palabra	Aptitud	Palabra	Aptitud	Palabra	Aptitud
rovhe	159	xahxoc	97	vozbah	81
fuho	134	foolzed	96	ridr	80
hsuvhee	129	idxeko	96	xibv	80
debo	127	ifhedo	96	udfihh	78
vexr	124	divu	95	vizaomf	76
foroxbo	123	corosi	95	ovlof	74
honii	123	kezku	94	gao	74
kexza	115	fihn	94	bavxic	73
rahn	112	koe	93	ofeniq	72
horizos	111	lurfee	93	gexfiq	71
dudorke	108	ozdocal	92	zluo	70
xeo	107	hima	88	fgiqe	70
exhegda	107	kamxurc	88	ohao	68
fesh	107	qiloh	85	eos	67
coim	101	ziruxx	85	mukdagc	67
reivusl	101	cmo	84	meivbu	67
rvivur	98	emudivz	82	bfe	66
zograrz	98	xue	82	hhi	66

<b>Palabra</b>	<b>Aptitud</b>	<b>Palabra</b>	<b>Aptitud</b>	<b>Palabra</b>	<b>Aptitud</b>
rucgic	66	sri	50	aqok	29
icqos	65	ibikuk	50	axd	28
dxa	64	bqacbi	49	ek	27
nhe	63	aofm	48	ulx	26
cqu	63	qla	43	ocg	26
his	62	mza	43	eml	24
eoq	61	gilc	43	eb	22
xud	60	ulile	42	emb	21
zeg	59	ulni	40	oa	20
zaaz	58	audz	39	afb	20
ehga	58	muv	39	alm	20
fxi	58	oud	37	ud	17
qizvi	57	olu	37	uo	12
buaq	56	egehg	37	aa	10
kas	56	uhod	37	do	10
gxe	55	gxi	34	re	10
kugs	55	ixacu	34	uu	9
afuoc	54	anez	33	iu	9
eqo	53	aga	33	be	9
ahoar	53	iqa	33	ko	9
ebbi	52	ex	32	no	9
fil	51	ox	32	ma	6
nuav	51	iuf	31	di	6

Tabla 3.5 Resultado de la ejecución del programa, un conjunto de palabras que han sido evaluadas

Una vez que se seleccionaron las palabras del conjunto solución para construir el nuevo lenguaje artificial, se asigna una palabra generada a cada palabra de la lista *swadesh*.

<b>Significado</b>	<b>Palabra</b>	<b>Aptitud</b>
Yo	em	40
Tú	faro	170
Nosotros	eivsov	120
Esto	fivc	129
Eso	docza	175
¿Quién?	neida	129
¿Qué?	dodax	169
Nosotros	sors	150
Todos	sanés	149
Muchos	eleuve	112
Uno	uhxo	123
Dos	vao	112
Grande	iorosr	116
Largo	beru	149
Pequeño	cosi	161
Mujer	ceqnu	147
Hombre	qokuo	147
Persona	nehvo	171
Pez	masbi	119
Ave	vdoa	126
Perro	onuqko	104
Casa	kizna	116
Árbol	evalso	125
Semilla	fuzd	114
Hoja	vee	114
Raíz	kreflo	133
Corteza	ceez	137
Piel	roboe	195
Carne	orsa	109
Sangre	xush	124
Hueso	keuve	138
Grasa	coudm	155
Huevo	rimoo	158
Cuerno	uolor	158
Cola	moba	139
Pluma	ovzuzk	109

<b>Significado</b>	<b>Palabra</b>	<b>Aptitud</b>
Pelo	urmeu	100
Cabeza	enhei	119
Oreja	navr	131
Ojo	vho	100
Nariz	ixaado	101
Boca	osafah	105
Diente	hulmevq	103
Lengua	voz	101
Uña	zenuv	137
Piel	fugbo	131
Rodilla	redamfi	96
Mano	cishih	111
Panza	anza	89
Cuello	duisx	114
Senos	lindee	112
Corazón	bara	147
Riñón	furbee	112
Beber	vohi	177
Comer	hodkeis	112
Morder	oaclind	90
Ver	kikaf	87
Escuchar	egivo	87
Saber	seberqe	121
Dormir	uuxk	82
Morir	aruok	90
Matar	qaezuzo	94
Nadar	eiqvovd	107
Volar	agzuido	86
Caminar	anae	84
Recostar	kuam	82
Mentir	muic	75
Sentarse	izbab	78
Pararse	ineavub	83
Dar	aboav	86
Decir	meh	63
Sol	kaomlan	79

Significado	Palabra	Aptitud	Significado	Palabra	Aptitud
Luna	xhuuxee	121	Rojo	ear	55
Estrella	ugfifm	80	Verde	kub	47
Agua	mek	60	Amarillo	lre	69
Lluvia	oxdum	93	Blanco	onn	35
Piedra	eugd	60	Negro	ox	40
Arena	ezag	66	Noche	xqedre	128
Tierra	hecu	175	Caliente	ogfuev	107
Nube	nuaggel	76	Frío	kgaxuve	108
Humo	nvoqe	94	Lleno	afgie	80
Fuego	zan	59	Nuevo	evoel	92
Ceniza	zsohofa	130	Bueno	fesle	164
Quemar	qam	48	Redondo	cixq	128
Camino	devi	150	Seco	qace	144
Montaña	qicoh	143	Nombre	daino	129

Tabla 3.6 Asignación de las palabras generadas a la lista swadesh

### 3.3 Segunda fase de desarrollo

En el procesamiento del lenguaje natural la ambigüedad juega un papel importante. Charnaik (1996) acuñó el término *treebank* para leer las reglas de producción de una oración. Su objetivo era poder alimentar a los analizadores sintácticos en una forma más extensa y reducir la ambigüedad del lenguaje. Al ser una gramática sin contexto, para una gramática de lenguaje natural, las reglas de producción pueden no ser siempre las propuestas.

さらば少年の日

Adiós, días de juventud.

Interjección – sustantivo – preposición – sustantivo

Días de juventud, adiós.

Sustantivo – preposición – sustantivo - Interjección

P -> S1 interj

S1 -> sus S2

S2 -> prep sus

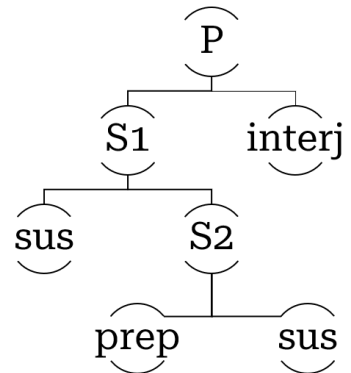


Figura 3.4 Análisis de una oración a partir de las reglas gramaticales de la nueva gramática

Este formato lo utilizó para la definición de la gramática del lenguaje. Para la obtención de la gramática he utilizado el algoritmo que Tomita (1984) utiliza para la obtención de analizadores sintácticos de una gramática, basado en el artículo de Lang (1974) para analizadores no determinísticos. El artículo menciona que para obtener la gramática se deben realizar los mismos pasos del analizador sintáctico LR. La diferencia en este caso es que como es una gramática sin contexto, lo que estamos buscando son las acciones, las reglas de producción, por lo que en el momento que las palabras se den sentido unas a otras, estas se reducen a una regla de producción. Y también menciona que, para una oración, se pueden generar múltiples y distintos árboles, dependiendo de cómo se introduzcan los elementos a la pila.

---

**Algoritmo 7** Analizador LR (Left to Right) Desplaza y Reduce

---

- 1: **para** cada simbolo de la expresión **hacer**
  - 2:     **si** individuo no se puede reducir **entonces**
  - 3:         Desplazar simbolo a la pila
  - 4:     **terminar si**
  - 5:     **si** individuo se puede reducir en una regla de producción **entonces**
  - 6:         Sacar elementos de la pila hasta que los requerimientos  
           de producción sean cumplidos
  - 7:         Desplazar a la pila la regla de producción
  - 8:     **terminar si**
  - 9: **terminar para**
- 

Reducir, si los elementos terminales tienen sentido, reducir a una regla de producción.

El lenguaje que se va a crear no tiene una gramática como tal, pero se tienen oraciones, por lo que se puede utilizar este método para deducir una posible gramática del lenguaje deseado.

Como parámetros decidí por oraciones en español, inglés, latín, japonés y chino. Estas oraciones fueron tomadas de diferentes contenidos: novelas, películas, series, manga, canciones y *tweets*. Frases que son importantes para mí y que quisiera fuera posible comunicarlas en la nueva lengua.

A partir de las frases analizadas, que se presentan a continuación, y de haber obtenido sus reglas de producción, sólo falta eliminar reglas duplicadas y reducir reglas que puedan generar los mismos resultados de otras. El conjunto final es la gramática del nuevo lenguaje.

### 3.3.1 Desarrollo de gramática

A continuación, se presenta el treebank. Una tabla con el análisis lingüístico de las oraciones, traduciéndolas si es necesario, el análisis gramatical, la propuesta de estructura para el nuevo

idioma, el reacomodo de los elementos para el nuevo idioma y las reglas de producción de éste último.

Frase original	Traducción	Análisis	Estructura en nuevo idioma	Análisis de nueva estructura	Reglas
さらば少年の 日	Adiós, días de juventud.	interjección sustantivo preposición sustantivo	Días de juventud, adiós.	sustantivo preposición sustantivo interjección	P -> S1 interj S1 -> sus S2 S2 -> prep sus P -> SNSN interj SNSN -> sus PRESN PRESN -> prep SN sus -> días   juventud prep -> de interj -> adiós
じゃあね、ま た来るよね！	Nos vemos, ¡regreso después!	interjección verbo adverbio	¡Después regreso, nos vemos!	adverbio verbo interjección	P -> S1 interj S1 -> adv S2 S2 -> verb adv -> después verb -> regreso interj -> nosvemos
¿Incluso si es una mala decisión?		adverbio conjunción condicional contabilizador de sustantivo adverbio sustantivo	¿Decisión mala incluso?	sustantivo adverbio adverbio	P -> S1 adv S1-> sus adv sus -> Decisión adv -> mala   incluso
I have crossed the horizon to find you.	He cruzado el horizonte para encontrarte.	pronombre verbo artículo sustantivo conjunción verbo	Cruzado horizonte encontrar tú.	verbo sustantivo verbo pronombre.	P -> S1 S2 S1 -> verb sus S2 -> verb pron verb -> curazo   encontrar sus -> horizonte prono -> tú

<p>No te preocupes, siempre tendremos el Internet.</p>		<p>adverbio pronombre verbo adverbio verbo artículo sustantivo</p>	<p>No preocupación tú, tener nosotros lugar internet.</p>	<p>adverbio verbo pronombre verbo pronombre sustantivo</p>	<p>P -&gt; S1 S2 S1 -&gt; S3 pron S3 -&gt; adv verb S2 -&gt; S4 S5 S4 -&gt; verb pro S5 -&gt; sus sus sus -&gt; lugar   internet adv -&gt; No verb -&gt; preparación   tener pron -&gt; tú   nosotros</p>
<p>Ya no hay emergencias, Luisa.</p>		<p>adverbio adverbio verbo sustantivo sustantivo</p>	<p>Emergencia no hay, Luisa.</p>	<p>sustantivo adverbio verbo sustantivo</p>	<p>P -&gt; S1 sus S1 -&gt; sus S2 S2 -&gt; adv verb sus -&gt; emergencia   Luisa adv -&gt; no verb -&gt; hay</p>
<p>A la única conclusión a la que llego es que no quiere nada de mí...</p>		<p>preposición artículo adjetivo sustantivo preposición artículo preposición artículo conjunción verbo verbo preposición adverbio verbo sustantivo preposición pronombre</p>	<p>Conclusión llegar yo no querer nada de mí.</p>	<p>sustantivo verbo pronombre adverbio verbo sustantivo preposición pronombre</p>	<p>P -&gt; S1 S2 S1 -&gt; sus verb S2 -&gt; pro S3 S3 -&gt; adv S4 S4-&gt; verb sus S5 S5 -&gt; prep pron sus -&gt; conclusión   nada verb -&gt; llegar   querer pron -&gt; yo   mí adv -&gt; no prep -&gt; de</p>

Te amo.		pronombre verbo	Todo mi amor para ti.	adjetivo pronombre sustantivo preposición pronombre	P -> S1 prep pron S1 -> adj S2 S2 -> pron sus adj -> todo pron -> mi   ti sus -> amor prep -> para
Never one... without the other.	Nunca uno... sin el otro.	adverbio pronombre preposición artículo adjetivo	Nunca uno... no otro sin.	adverbio pronombre adverbio adjetivo preposición	P -> S1 S2 S1 -> adv pron S2 -> S3 prep S3 -> adv prep adv -> Nunca   no pron -> uno prep -> sin
Por mucho que me esfuerce a todas horas, el deseo salta al revés.		preposición adverbio conjunción artículo verbo preposición adverbio sustantivo artículo sustantivo verbo adverbio	Demasiado esfuerzo, salta deseo dirección atrás.	adjetivo sustantivo verbo sustantivo sustantivo adverbio	P -> S1 S2 S1 -> S3 S2 S2 -> S4 S5 S4 -> verb sus S5 -> sus adv S3 -> adj sus adj -> demasiado sus -> esfuerzo   deseo   dirección verb -> salta   adv
I have tried to break off this bunch and hand it you; but whether there is substance or	He tratado de coger este racimo de mi vida y ofrecéroslo, pero si hay substancia o verdad en ella, no lo sé.	pronombre verbo preposición verbo determinante objeto preposición determinante	Intento recoger racimo vida misma y ofrecer para ti, no saber existe substancia o realidad en racimo, no sé.	adverbio verbo sustantivo sustantivo - adjetivo conjunción verbo preposición pronombre adverbio verbo	P -> S1 S5 S1 -> sus sus S7 -> S1 adj S8 -> prep pron S2 -> sus conj S4 -> prep sus S5 -> adv verb S6 -> S5 verb

truth in it I do not know.		sustantivo conjunción verbo conjunción condicional verbo sustantivo conjunción sustantivo preposición artículo adverbio pronombre verbo		verbo sustantivo conjunción sustantivo preposición sustantivo adverbio verbo	S3 -> S2 sus S9 -> S3 S4 S13 -> S6 S9 S10 -> S5 S7 S12 -> S10 conj S11 -> verb S8 S14 -> S12 S11 S15 -> S14 S13 sus -> racimo   vida   sustancia   realidad adv -> no   intento verb -> sé   recoger   ofrecer   saber   existe adj -> misma conj -> y   o prep -> para   en pron -> ti
I read somewhere the trick is to keep breathing...	En alguna parte leí que el truco es seguir respirando...	preposición adverbio adjetivo sustantivo verbo conjunción artículo verbo verbo verbo verbo	Leí lugar truco seguir respirando.	verbo sustantivo sustantivo verbo verbo	P -> S1 S2 S2 -> verb verb S3-> verb sus S1 -> S3 sus sus -> lugar   truco verb -> leí   seguir   respirando
There... Now I've ruined everything.	Ya... Lo he arruinado todo.	adverbio adjetivo pronombre verbo	Arruinado todo... ya.	verbo adjetivo adverbio	P -> verb S1 ADJADV -> adj adv adv -> todo verb -> arruinado adj -> adverbio

		verbo adjetivo			
No me digas eso, Raimunda, que me pongo a llorar... Y los fantasmas no lloran.		adverbio pronombre verbo pronombre sustantivo conjunción pronombre verbo preposición verbo conjunción artículo sustantivo adverbio verbo	No decir eso, Raimunda, llorar yo comienzo. No lloran fantasmas.	adverbio verbo pronombre sustantivo verbo pronombre sustantivo adverbio verbo sustantivo	P -> S1 S8 S3 -> adv verb S2 -> S3 pron S1 -> S2 sus S4 -> pron sus S7 -> verb S4 S5 -> verb sus S6 -> adv S5 S8 -> S7 S6 adv -> no verb -> decir   llorar   lloran pron -> eso   yo sus -> Raimunda   comienzo   fantasmas
The majority of our sufferings in life result from the dreams of others.	La mayoría de nuestros sufrimientos en la vida son el resultado de los sueños de otros	artículo sustantivo preposición pronombre sustantivo preposición artículo sustantivo verbo artículo sustantivo preposición artículo sustantivo preposición sustantivo	Sufrimientos de uno mismo al vivir son resultado sueños otros.	sustantivo preposición pronombre adjetivo contracción verbo verbo sustantivo sustantivo adjetivo	P -> S8 S7 S1 -> sus prep S2 -> pron adj S3 -> cont verb S4 -> verb sus S5 -> sus adj S6-> S1 S2 S7-> S4 S5 S8-> S6 S3 sus -> sufrimientos   resultados   sueños prep -> de pron -> uno adj -> mismo   otros cont -> al

					verb -> vivir   son
すすめいとし い風よ	Avanza, mi querido viento.	adverbio adverbio adjetivo verbo sustantivo	Adelante adelante, querido viento.	adverbio adverbio adjetivo sustantivo	P -> S1 S2 S1 -> adv adv S2-> adj sus adv -> adelante adj -> querido sus -> viento
どうか忘れな いで…ボクは キミを心から 愛している と。	Por favor, no me olvides... Desde el fondo de mi corazón te amo.	preposición sustantivo adverbio pronombre verbo preposición artículo adjetivo preposición adjetivo sustantivo pronombre verbo	No olvidar mí... fondo corazón mío todo mi amor para ti.	adverbio verbo pronombre sustantivo sustantivo pronombre adjetivo pronombre sustantivo preposición pronombre	P -> S1 S9 S1 -> adv verb S2 -> sus sus S3 -> pron adj S4 -> pron sus S5 -> prep pron S6 -> pron S2 S7 -> S4 S5 S8 -> S6 S3 S9 -> S8 S7 adv -> no verb -> olvidar pron -> mi   mío   ti sus -> fondo   corazón   amor adj -> todo prep -> para
A great fire burns within me, but no one stops to warm themselves at it, and passers- by only see a wisp of smoke.	Un gran fuego arde dentro de mí, pero nadie se detiene a sentir su calor, los transeúntes sólo ven una voluta de humo.	artículo intensificador sustantivo verbo adverbio preposición artículo pronombre	Gran Fuego arde dentro mío, detenerse nadie calor sentir, personas pasar a distancia y ver nube humo.	adjetivo sustantivo verbo adverbio adjetivo verbo pronombre sustantivo verbo	P -> S11 S5 S1 -> adj sus S8 -> adv adv S6 -> verb S1 S9 -> S8 S6 S7 -> verb pron S5 -> sus verb S10-> S7 S5

		conjunción sustantivo pronombre verbo preposición verbo adjetivo sustantivo articulo sustantivo adverbio verbo adjetivo sustantivo preposición sustantivo		sustantivo verbo preposición sustantivo conjunción verbo sustantivo sustantivo	S4 -> prep sus S2 -> sus sus S3 -> verb S2 S5 -> conj S3 S11 -> S9 S10 adj -> gran   dentro sus -> fuego   calor   persona   distancia   nube   humo adv -> de   mío verb -> arde   detenerse   sentir   pasar   ver pron -> nadie prep -> a conj -> y
We are such dreary liars. Our one idea is to lie to ourselves.	Somos solo tristes mentirosos, nuestra única idea es mentirnos a nosotros mismos.	verbo adjetivo adjetivo adjetivo pronombre adjetivo sustantivo verbo verbo preposición pronombre adjetivo	Mentirosos tristes somos, idea única nuestra es nosotros mentir mismos.	adjetivo adjetivo verbo sustantivo adjetivo adjetivo verbo pronombre verbo adjetivo	P -> S2 S8 S1 -> sus sus S2 -> S1 verb S3 -> sus adj S4 -> S3 adj S5 -> S4 verb S6 -> verb adj S7 -> pron S6 S8 -> S5 S7 sus -> idea verb -> somos   es   mentir adj -> mentirosos   tristes   única   nuestra   mismos pron -> nosotros
If you intend to fall into this	Si tu intención es caer en este	conjunción adjetivo	Intención tuya caer en este	sustantivo adjetivo verbo	P -> S6 S3 S1 -> sus adj S4 -> S1 verb

<p>Abyss, I'll fall with you.</p>	<p>abismo, caeré contigo.</p>	<p>sustantivo verbo verbo preposición sustantivo sustantivo verbo pronombre</p>	<p>abismo, contigo caeré</p>	<p>preposición sustantivo sustantivo pronombre verbo</p>	<p>S2 -&gt; prep adj S5 -&gt; S2 sus S3 -&gt; pron verb S6 -&gt; S4 S5 sus -&gt; intención   este   contigo adj -&gt; tuya verb -&gt; caer   caeré prep -&gt; en pron -&gt; contigo</p>
<p>En todos esos cielos brilla igual nuestra luna llena y tú sigues siendo la mejor.</p>		<p>preposición adjetivo pronombre sustantivo verbo adjetivo pronombre sustantivo adjetivo conjunción artículo adverbio verbo artículo adjetivo</p>	<p>Todo eso brillar en cielo, llena luna nuestra y tú la mejor seguir siendo</p>	<p>adjetivo sustantivo verbo preposición sustantivo adjetivo sustantivo adjetivo conjunción pronombre artículo adjetivo verbo verbo</p>	<p>P -&gt; S10 S11 S1 -&gt; adj sus S2 -&gt; prep sus S8 -&gt; verb S2 S3 -&gt; adj sus S7 -&gt; S1 adj S4 -&gt; art adj S5 -&gt; verb verb S6 -&gt; conj pron S9 -&gt; S4 S5 S10 -&gt; S3 S8 S11-&gt; S7 S6 S9 adj -&gt; todo   llena   nuestra   mejor sus -&gt; eso   cielo   luna verb -&gt; brillar   seguir   siendo prep -&gt; en conj -&gt; y pron -&gt; tú art -&gt; la</p>
<p>De mirarlo desde fuera con una mirada que él había creído reservada para</p>		<p>preposición verbo preposición adverbio preposición</p>	<p>Desde fuera mirar con mirada única creer él reserva</p>	<p>preposición adverbio verbo preposición sustantivo</p>	<p>P -&gt; S10 S11 S1 -&gt; prep adv S2 -&gt; verb prep S4 -&gt; sus adj S3 -&gt; verb pron</p>

siempre a los recuerdos.		adjetivo sustantivo conjunción pronombre verbo verbo adjetivo preposición adverbio preposición artículo sustantivo	siempre para el recuerdo.	adjetivo verbo pronombre verbo adverbio preposición artículo sustantivo	S5 -> verb adv S6 -> art sus S7 -> prep S6 S9 -> S3 S5 S8 -> S2 S4 S10 -> S1 S8 S11 -> S9 S7 adj -> única sus -> mirada   recuerdo verb -> mirar   reserva   adv -> fuera   creer   siempre prep -> desde   con   para pron -> él art -> el
Despite your violent behavior, the only thing you've managed to break so far is my heart.	A pesar de tu comportamiento violento, la única cosa que has logrado es romper mi corazón	preposición verbo preposición artículo sustantivo adjetivo artículo adjetivo sustantivo conjunción verbo verbo conjunción verbo adjetivo sustantivo	a pesar de tu comportamiento violento, la única cosa que has logrado es romper mi corazón.	preposición preposición sustantivo pronombre adjetivo sustantivo adjetivo sustantivo verbo sustantivo verbo adjetivo	P -> S9 S10 S1 -> prep rep S2 -> sus pron S3 -> S2 adj S9 -> S1 S3 S4 -> sus adj S5 -> S4 sus S6 -> S5 vreb S7 -> sus verb S8 -> S7 adj S10 -> S6 S8 adj -> violento   única   mío sus -> comportamiento   cosa   logro   corazón verb -> es   romper prep -> a   pesar pron -> tuyo

Love is by definition an unmerited gift; being loved without meriting it is the very proof of real love.	El amor por definición es un regalo que no requiere méritos; Ser amado sin ameritarlo es la prueba del amor verdadero.	artículo sustantivo preposicion sustantivo verbo pronombre sustantivo pronombre adverbio verbo sustantivo verbo verbo preposicion ameritarlo verbo artículo sustantivo preposicion sustantivo adjetivo	Por definición amor es regalo que méritos no requerir; Amado ameritar sin, prueba verdadero amor.	preposición sustantivo sustantivo verbo sustantivo pronombre sustantivo adverbio verbo sustantivo verbo preposición sustantivo preposición sustantivo	P -> S12 S9 S1 -> prep sus S2 -> sus verb S3 -> S2 sus S10 -> S1 S3 S4 -> adv verbo S5 -> sus S4 S11 -> pron S5 S6 -> adj sus S7 -> sus S6 S8 -> prep S7 S9 -> S2 S8 S12 -> S10 S11 adj -> verdadero adv -> no sus -> definición   amor   regalo   meritos   amado   prueba verb -> es   requerir   amerita prep -> Por   sin pron -> que
To love at all is to be vulnerable.	Tan sólo amar es ser vulnerable.	adverbio adjetivo verbo conjunción verbo adjetivo	Amar sólo vulnerable ser	verbo adjetivo adjetivo verbo	P -> S2 verb S1 -> adj adj S2 -> verb S1 adj -> sólo   vulnerable verb -> amar   ser
No importa si no te veo hoy o mañana, te veré después sin problemas.		adverbio verbo conjunción adverbio pronombre verbo	Importar no ver a ti hoy o mañana, veré a ti después, problemas no.	verbo adverbio verbo preposición pronombre adverbio	P -> S10 S7 S1 -> adv verb S2 -> prep pron S3 -> adv conj adv S4 -> S1 S2 S5 -> S4 S3

		sustantivo conjunción sustantivo pronombre verbo adverbio preposición sustantivo		conjunción adverbio verbo preposición pronombre adverbio sustantivo adverbio	S9 -> verb S5 S6 -> verb S2 S8 -> S6 adv S7 -> sus adv S10 -> S9 S8 sus -> problemas verb -> importar   ver   veré   después adv -> no   hoy   mañana prep -> a   conj -> o pron -> ti
Somnias memorias vienen y van como sueños mil.		adverbio sustantivo verbo conjunción verbo adverbio sustantivo cuantificador	Memorias mías venir ir sueños mil cómo	sustantivo adverbio verbo verbo sustantivo adjetivo adverbio	P -> S4 S5 S1 -> verb verb S2 -> S1 sus S3 -> S2 adj S5 -> S3 adv S4 -> sus adv adj -> mil sus -> memorias   sueños adv -> mías   cómo verb -> venir   ir
そんな顔をし ないで。。。 また会える よ。	No hagas esa cara... Nos volveremos a ver.	adverbio verbo pronombre sustantivo pronombre verbo conjunción verbo.	Tú cara no hacer... volver a ver nosotros	pronombre sustantivo adverbio verbo verbo preposición verbo pronombre	P -> S5 S6 S1 -> pron sus S2 -> adv verb S3 -> verb prep S4 -> verb pron S5 -> S1 S2 S6 -> S3 S4 sus -> cara adv -> no verb -> hacer   volver   ver prep -> a

					pron -> tú   nosotros
Imáname: no puedo existir si no me imaginas.		verbo adverbio verbo verbo conjunción adverbio pronombre verbo.	Imaginar tú yo: existir no yo si imaginar yo no.	verbo pronombre pronombre verbo adverbio pronombre conjunción verbo pronombre adverbio.	P -> S7 S6 S1 -> pron pron S2 -> S1 verb S7 -> verb S2 S3 -> adv pron S4 -> pron adv S5 -> verb S4 S6 -> S3 S5 verb -> imaginar   existir adv -> no conj -> si pron -> tú   yo
ここには神様が いるの	¿Dios existira?	sustantivo verbo	Existir dios?	verbo sustantivo	P -> verb sus sus -> dios verb -> existir
Mira esa piedra cómo ya no se para...		verbo pronombre sustantivo adverbio adverbio adverbio verbo verbo	Mirar piedra allá ya no detener...	verbo sustantivo adverbio adverbio adverbio verbo	P -> S1 S4 S1 -> verb sus S2 -> adv adv S3 -> adv S2 S4 -> S3 verb sus -> piedra verb -> mirar   detener adv -> allá   ya   no
I don't believe that.	No creo eso.	adverbio verbo pronombre	Eso no creo.	adjetivo adverbio verbo	P -> adj S1 S1 -> adv verb adj -> eso verb -> creo adv -> no

It's been a long, long time since I've memorized your face...	Ha pasado mucho, mucho tiempo desde que memoricé tu cara...	verbo adjetivo adjetivo adjetivo sustantivo preposición conjunción verbo adjetivo sustantivo	Mucho tiempo pasa, mucho tiempo desde tu cara memorizar	adjetivo sustantivo verbo adjetivo sustantivo preposición adjetivo sustantivo verbo	P -> S4 S3 S1 -> adj sus S2 -> S1 verb S3 -> prep S2 S4 -> S2 S1 adj -> mucho   tu sus -> tiempo   cara verb -> pasa   memorizar prep -> desde
Cuando bailo contigo, no me preocupo más.		adverbio verbo contigo adverbio pronombre verbo adverbio	Bailo contigo momento, preocupo yo no más.	verbo pronombre sustantivo verbo pronombre adverbio adverbio	P -> S2 S4 S1 -> verb pron S2 -> S1 sus S3 -> adv adv S4 -> S1 S3 sus -> momento verb -> bailo   preocupo adv -> no   más pron -> contigo   yo
Todo lo que queda es el juego de sueños y el extraño recuerdo		adjetivo artículo conjunción verbo verbo artículo sustantivo preposición sustantivo conjunción artículo adjetivo sustantivo.	Solo queda sueño de juegos y recuerdo extraño	adjetivo verbo sustantivo preposición sustantivo conjunción sustantivo adjetivo	P -> S1 S5 S1 -> adj verb S2 -> prep sus S4 -> sus S2 S3 -> sus adj S5 -> S4 conj S3 adj -> sólo   extraño sus -> sueño   juegos   recuerdo verb -> queda prep -> de conj -> y

Las mejores historias no nos relajan, nos desafían.		sustantivo adjetivo adverbio verbo pronombre verbo pronombre	Historias mejores no rejarlar nosotros, desafiar nosotros.	sustantivo adjetivo adverbio verbo pronombre verbo pronombre	P -> S5 S4 S1 -> sus adj S2 -> adv verb S3 -> S2 pron S4 -> verb pron S5 -> S1 S3 adj -> mejores sus -> historias adv -> no verb -> relajar   desafiar pron -> nosotros
要找不到才行 啊! 我这个人, 什么都没有, 有的是时间, 我总得找点事情来打发时间	Si encuentro a la persona indicada... Un hombre como yo no tiene nada, excepto tiempo libre.	conjunción verbo preposición artículo sustantivo adjetivo adjetivo sustantivo adverbio pronombre adverbio verbo sustantivo conjunción sustantivo adjetivo	Encontrar yo indicada persona... Hombre parecido yo nada tener, tiempo libre solo	verbo pronombre adjetivo sustantivo sustantivo adjetivo pronombre sustantivo verbo sustantivo adjetivo adjetivo	P -> S8 S9 S1 -> adj sus S2 -> pron S1 S8 -> verb S2 S3 -> sus adj S4 -> sus verb S5 -> pron S4 S7 -> S3 S5 S6 -> S3 adj S9 -> S7 S6 adj -> indicada   parecido   libre   sólo sus -> persona   hombre   nada   tiempo verb -> encontraba   tener pron -> yo

Tabla 3.7 Análisis de cincuenta oraciones en diferentes idiomas a partir de las reglas de la nueva gramática, y las reglas de producción generadas a partir de los análisis.

### 3.3.2 Reducción de las reglas de producción final

Al reducir las reglas de producción de las frases seleccionadas se obtiene un conjunto manejable. En la Tabla 3.7 se muestran cinco niveles de producción. Cada nivel contiene el nivel inferior, siendo el Nivel 0 aquél con las reglas base.

Esto quiere decir que muchos tipos de frases cortas y sólo pocas frases largas, casos muy específicos, se pueden generar en el nuevo lenguaje.

Nivel 0	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5
S1 -> verb sus	K1 -> sus S2	M1 -> K7 K4	O1 -> pron M3	L1 -> S14 O1	Z1 -> L2 S3
S2 -> prep sus	K2 -> S7 adv	M2 -> K8 sus	O2 -> M6 K12	L2 -> O4 O3	Z2 -> L7 S7
S3 -> adv verb	K3 -> S1 S4	M3 -> adv K9	O3 -> K18 M9	L3 -> S3 O7	Z3 -> M40 L10
S4 -> verb pron	K4 -> S4 S8	M4 -> K10 S6	O4 -> M8 conj K20	L4 -> K27 O8	
S5 -> pron sus	K5 -> S1 interj	M5 -> S9 K11	O5 -> M11 M10	L5 -> M22 O12	
S6 -> prep pron	K6 -> S3 interj	M6 -> S1 K12	O6 -> M12 K23	L6 -> O14 O13	
S7 -> sus adv	K7 -> S3 pron	M7 -> K13 S12	O7 -> M13 K26	L7 -> O15 M27	
S8 -> sus sus	K8 -> sus S3	M8 -> S3 K17	O8 -> M14 K29	L8 -> S7 O16	
S9 -> adv pron	K9 -> S1 S6	M9 -> K19 S10	O9 -> M16 S24	L9 -> M38 O22	
S10 -> prep sus	K10 -> adj S5	M10 -> K22 K21	O10 -> M17 M18	L10 -> O23 M41	
S11 -> adj sus	K11 -> S9 prep	M11 -> K7 sus	O11 -> M19 M20		
S12 -> verb verb	K12 -> S1 S7	M12 -> K24 S20	O12 -> M21 K43		
S13 -> adv adv	K13 -> S1 sus	M13 -> K25 S18	O13 -> S14 M23		
S14 -> sus verb	K14 -> verb S19	M14 -> K28 verb	O14 -> M25 M24		
S15 -> adv prep	K15 -> S13 S11	M15 -> K33 verb	O15 -> verb M26		
S16 -> sus conj	K16 -> verb S8	M16 -> K32 K33	O16 -> M28 adv		
S17 -> sus prep	K17 -> S8 adj	M17 -> S11 K37	O17 -> M30 M31		
S18 -> pron adj	K18 -> S3 verb	M18 -> K34 S26 K35	O18 -> S1 M32		
S19 -> adj adv	K19 -> S16 sus	M19 -> S27 K40	O19 -> M33 M34		
S20 -> cont verb	K20 -> verb S6	M20 -> K39 K38	O20 -> S36 M36		
S21 -> sus adj	K21 -> adv S1	M21 -> K42 verb	O21 -> M37 S4		
S22 -> verb adj	K22 -> verb S5	M22 -> S31 K41	O22 -> M39 K28		
S23 -> prep adj	K23 -> S1 S21	M23 -> prep K45	O23 -> K52 M41		
S24 -> pron verb	K24 -> S17 S18	M24 -> pron K8			
S25 -> art adj	K25 -> pron S8	M25 -> S10 K44			

S26 -> conj pron	K26 -> S5 S6	M26 -> K46 D1
S27 -> prep adv	K27 -> S8 verb	M27 -> K20 adv
S28 -> verb prep	K28 -> S21 adj	M28 -> K47 adj
S29 -> verb adv	K29 -> pron S22	M29 -> K49 K48
S30 -> art sus	K30 -> adj S3	M30 -> verb K50
S31 -> prep prep	K31 -> verb S33	M31 -> S9 K51
S32 -> sus pron	K32 -> S21 verb	M32 -> K52 verb
S33 -> adj adj	K33 -> S23 sus	M33 -> prep K53
S34 -> pron pron	K34 -> S11 adj	M34 -> K53 S11
S35 -> pron adv	K35 -> S25 S12	M35 -> K54 K55
S36 -> adv verb	K37 -> verb S10	M36 -> K56 conj S21
D1 -> adv conj adv	K38 -> prep S30	M37 -> S21 K7
	K39 -> S4 S29	M38 -> verb K57
	K40 -> S28 S21	M39 -> S21 K58
	K41 -> S32 adj	M40 -> S11 K60
	K42 -> S21 sus	M41 -> conj K16
	K43 -> S14 adj	
	K44 -> S14 sus	
	K45 -> sus S11	
	K46 -> S3 S6	
	K47 -> S12 sus	
	K48 -> S5 S3	
	K49 -> S28 S4	
	K50 -> S34 verb	
	K51 -> verb S35	
	K52 -> adv S13	
	K53 -> S11 verb	
	K54 -> S4 sus	
	K55 -> S4 S13	
	K56 -> sus S10	
	K57 -> pron S11	
	K58 -> pron S14	
	K59 -> verb S11	
	K60 -> verb S13	
	K61 -> S4 S14	
	K62 -> S14 S10	

Tabla 3.8 Tabla final de reglas de producción después de la reducción

Con este segmento de gramática es posible generar nuevas oraciones, por ejemplo:

Reglas de producción utilizadas	Frase reducida en palabras	Frase en el nuevo idioma	Frase original
M2 -> K8 sus K8 -> sus S3 S3 -> adv verb P -> sus adv verb sus	camino nuevo  camina persona	devi evoel anae  nehvo	La persona camina  por nuevo camino.
M30 -> verb K50 K50 -> S34 verb S34 -> pron pron P -> verb pron pron verb	mentir yo tú saber	muic em faro  seberqe	Tú siempre sabes  cuando miento.
O4 -> M8 conj K20 M8 -> S3 K17 K20 -> verb S6 K17 -> S8 adj S3 -> adv verb S6 -> prep pron S8 -> sus sus P -> adv verb sus sus adj conj verb prep pron	seco decir tú ojo  rojo y caminar  contra yo	qace meh faro vho  ear di anae zenuv  em	Me dices de forma  seca que tienes el  ojo rojo y te alejas  de mí.
M37 -> S21 K7 S21 -> sus adj K7 -> S3 pron S3 -> adv verb	hombres grandes  rápido derrotado yo	qokuo iorosr onn  qaezuzo em	Derroté rápidamente  a los hombres  grandes

P -> sus adj adv verb pron			
M15 -> K33 verb K33 -> S23 sus S23 -> prep adj P -> prep adj sus verb	Sobre grande raíz  sentar	(em) iorosr kreflo  izbab	Me senté sobre la  gran raíz

*Tabla 3.9 Ejemplos de oraciones en el nuevo lenguaje, después de haber obtenido las reglas de producción y el conjunto de palabras*

## 4 Daino gazezuko

Un grito desgarrador resuena desde el fondo del abismo. Garalla, sostenida de un risco, reta la oscuridad con una mirada vacía, casi molesta. Desde que comenzó el descenso su determinación había comenzado a amainar. Temía no estar tomando la mejor decisión; Pero, que el mismo terror se burlara de su patético intento, le parecía irritante.

—Ya sé que no voy a regresar —se dijo para sus adentros con una resignación pesada, densa y continuó descendiendo.

Pensó en su familia, lo lejos que estaba, lo poco que había podido ayudarla.

\* \* \*

—*Osafah*. —su abuela le había dicho una tarde—. Pasando el *cosi qicoh* al norte, se encuentra una cueva a la que nunca debes entrar. Mi niña, mi *coceqni* querida, prométeme que nunca irás para allá. No hay nada allá para ti. *¿Bara meh?*

—*Bara meh*, abuelita, *bara meh*. Los *seberqe* ya me explicaron por qué estamos aquí, quiénes somos y por qué vivimos tan cerca del *Osafah*.

—Lo sé, —una cálida pero triste sonrisa asomó en su rostro— ya no eres una niña. Ya eres grande, sí. Pero he vivido cosas que intento constantemente olvidar... Recuerdos que, cuando te veo, vuelvo a revivir. Garalla, yo temo que tengas que pasar por lo mismo.

\* \* \*

Garalla no pudo procesar la muerte de su abuela. Acompañó desde cierta distancia la procesión, pero en ningún momento quiso entablar conversación con nadie. En su comunidad la muerte es una transición importante, el relevo de una vida a otra.

Incluso Garalla lo cree, aunque no tenía las energías para demostrarlo, no hubiera podido. Se sintió ajena a la celebración. Los odiaba por estar celebrando.

¿Por qué tenía que aceptar así, sin pensarlo mucho, su muerte?

Ella le había dicho que nunca la abandonaría. ¿Por qué entonces se había ido? ¿Por qué lo había hecho tan pronto? Garalla había decidido que descendería el *Osafah* para volverla a ver. Arriesgaba su posición en la aldea al entrar en un área prohibida sólo para buscar a su abuela.

Sin ella dejó de ser alguien, todo había perdido importancia. Ya no había nadie más que le dijera que era buena, que era inteligente, que haría todos sus sueños realidad. Su lugar en la vida de un día para otro desapareció.

En un último intento de salvarse: la fue a buscar.

\* \* \*

Nunca había escalado. Le resultaba casi imposible sin luz, pero tenía que hacerlo. Era un proceso lento colocar manos y pies en una posición estable y segura. Pero estaba llena de determinación. Lo más difícil, llegar a *Osafah*, ya lo había logrado. El descenso de pronto terminó. No. Al intentar colocar uno de sus pies, no encontró el apoyo que buscaba. Perdió el equilibrio y el agarre en sus manos. Cayó varios metros en caída libre. Con un sonido seco, su cuerpo golpeó una cama de piedras, pero no siguió cayendo. Percibió el olor húmedo de un acueducto, el calor de agua caliente, el olor a azufre. Antes de perder el conocimiento comenzó a oír el sonido de agua correr, de agua caer desde algún lugar alto. Sus brazos y piernas sintieron el peso completo del ejercicio, de la caída. En la oscuridad se desvaneció.

\* \* \*

Tiempo después de que su abuela muriera, Garalla comenzó a tener visiones, sueños, pesadillas. La veía entrar en la cocina llenando el cántaro de agua, en su cuarto leyendo o platicando con su madre. Pero era como una sombra, algo que realmente no podía ver. Siempre en la periferia, en el límite de la que realmente existía.

Cuando se lo platicó a su madre, recibió un regaño como respuesta:

—Garalla, deja a los *aruonev* en paz. Puedo entender que estés triste por su partida, pero tú sabes lo importantes que son para que mi madre descanse.

—Lo sé, pero... ¡quiero volver a verla! Ella me contó que podía hacerlo. Si los *auronev* siguen cerca, significa que todavía hay tiempo para volver a verla.

—¡Recuerda los momentos bonitos que tuviste con ella! Eso... debería ser suficiente.

—¡No lo es! No me estás entendiendo. De nada sirve lo que te estoy contando.

—¡*Faro anae kizna!* —y cuando Garalla ya había salido, gritó tras de ella— ¡Y deja de buscarlos!

\* \* \*

La despertó un dolor en sus costillas. Intentó levantarse, pero no lo logró, tan sólo pudo girar un poco. Siguió acostada, sin poder reconocer nada dentro de la oscuridad. Su respiración, además de errática por el dolor, era lenta por la concentración alta de gases y vapores en el aire. Notó que la luz que se filtraba por la entrada de la cueva cuando inició su descenso ya no era visible. La luz de la luna, con la que había contado para su descenso, no existía en ese otro mundo. Uno a uno, iba perdiendo los sentidos.

—*Ceeznee*, — musitó con un dolor —. Ya no me falta mucho, pero me voy a tardar un poco más, espérame... —se levantó con esfuerzo y con mucho tiento volvió a ponerse en marcha.

\* \* \*

Cuando Garalla era una niña, su abuelita le contaba cuentos antes de dormir. La mayoría de ellos estaban cargados de tradición. Los contaba con tanto gusto que prefería escucharla a ella antes que a los *seberqe* en sus clases de historia. Los hacía tan interesantes que quien estuviera cerca siempre se acercaba un rato a escucharlos. Cada vez que contaba alguna leyenda conocida, su abuela le daba un toque personal que la mejoraba, y todos querían saber cuál sería el nuevo giro que le daría.

Una noche, cuando tenía diez años, su abuelita le dijo que ese día no le contaría un cuento. —Necesito que hablemos de algo importante. Algo que necesitas aprender. Ya casi eres grande.

Garalla la miró con anticipación. Como si le fuera a enseñar un truco de magia. Su abuela le contó que aprendería a ver a sus ancestros, platicar con ellos, conseguir consejos.

—*Vho kikaf ox, barek. Kxuve, sebenhei dodaino xhuuxee.* —Garalla cerró los ojos emocionada de lo que podría ver, una sonrisa se le escapó.

—*Nvoque kikaf, Ceeznee.* —y abrió los ojos.

—*Kxuven barenvo, fivc ioroino. Kaomxee em devi aurok...*—su abuela le regresó la sonrisa.

—*Beru kaomxee faro hecuz faro iosor egivaf.* Ahora, lo que tienes que hacer...

\* \* \*

Garalla avanzó por horas y horas por un túnel que no mostraba tener final. Poco a poco se volvía más estrecho y los gases cada vez más densos. Estaba tan cansada... Dentro de la oscuridad lágrimas comenzaron a acumularse y caer. Por más que intentaba acordarse de los cuentos que le había contado, no lograba acordarse de ninguno.

Ojalá su abuela le hubiera contado más cuentos...

—*Ceeznee...* —pero nadie, ni siquiera el eco que pudiera haber ocurrido, contestó.

La realidad era esa: ya no escucharía su voz, ya no la abrazaría, ya no estaría ahí para ella.

Por más que quisiera, ya no podría volver a verla.

—¡Ya no puedo seguir! ¡Perdóname! —Garalla dio dos, tres, pasos antes de caer de rodillas y llorar desoladamente. Le faltaba el oxígeno.

Una voz resonó lejos, pero no muy lejos.

—Mi *coceqni...*

En la distancia una luz gradualmente cayó desde arriba revelando un frondoso árbol. En la sombra del árbol estaba su abuelita recostada, dormida, respirando tranquilamente. Garalla se acercó y se sentó junto a ella, se acurrucó en su hombro. Estaba cansada, demasiado cansada, no podía más. Además de que estar allá abajo cada minuto se hacía más y más difícil.

—Por fin... —su voz ya no tenía fuerza. Las palabras se escapaban de sus labios junto a sus exhalaciones—. *Em ioxo kafuh. ¿evomlan emfarova uuxk?* —se acomodó un poco para voltear a verla a la cara. Irradiaba placidez.

Ya no estaban en una cueva. Es una tarde de verano, el sol en el horizonte calienta lo suficiente para no tener frío. El viento mece el pasto a su alrededor. Los pájaros planean a lo lejos, de árbol a árbol. El lugar está inundado de paz.

El cansancio era infinito. Sus párpados cada vez eran más pesados. Pero antes de que pudiera dormir le habló su abuela.

—Garalla. Garalla... —La sacudió de un hombro—. ¡Garalla! No te puedes dormir aquí. ¡No puedes estar aquí! Te dije que nunca vinieras a *Osafah*. ¡Vete! Anda, levántate y vete ya.

—P-Pero, ¡vine por ti! No me digas que ya... Yo quiero...

—Garalla... Yo ya no puedo estar contigo. Sé que ya no puedo platicar contigo, que ya no puedes escuchar mi voz, pero... Hay mucha gente que te quiere y te querrá. Ahí está tú mamá, tus amigas. Incluso me atrevo a decir que tienes muchos amigos en el pueblo. Escúchalas a ellas. —Su abuelita le acarició la mejilla con mucho amor—. Pide ayuda, vive, Garalla, seguro habrá algo que puedas hacer.

—Lo voy a intentar... pero si tan solo tú —pero ya no pudo terminar de hablar, se había vuelto a desvanecer.

\* \* \*

En la aldea ya se movilizaban. No mucho después de que desapareció Garalla, ya la habían comenzado a buscar. A lo largo del primer día que desapareció la buscaron en la aldea, en los alrededores, en la laguna cercana y en la colina de *Kuntred*, pero no lograban encontrar siquiera un rastro de a dónde había ido. Al inicio del segundo día decidieron que, aunque prohibido, tenían que entrar en *Osafah*. Asustada, su madre ya había pasado de pensar en regañarla. Sólo quería que Garalla regresara sana y salva. Todo estaría bien, todo se le perdonaría. Sólo quería que regresara.

\* \* \*

Despertó en su cuarto y lo único que pudo hacer fue llorar en la cama, adolorida, sin entender realmente cómo es que estaba otra vez en su cama. En ese momento fue que recordó que su abuelita le había dicho: que nunca fuera a *Osafah*, que ahí no había nada para ella y nunca lo habría. Una sonrisa afloró en su cansado rostro.

—Otra vez se me olvidó algo que me habías dicho... Perdón.

Garalla se sentó en la orilla de la cama. Respiró profundo, tomó todo el aire que sus pulmones pudieran guardar y exhaló lentamente. Con un poco de dificultad, puso sus pies descalzos en el suelo. La tierra estaba fría, fresca.

Garalla dio un paso adelante y luego otro...

\* \* \*

Una noche Garalla y su abuela discutían sobre su cama. No estaba de humor para escuchar una historia repetida.

—Ese ya me lo sé. Cuéntame otro.

Su abuela se puso a pensar... pero por alguna razón no recordaba otra historia.

—El único que me se ahorita es el de Ali Babá y los cuarenta ladrones... ¿ya te lo sabes?

—Abuelitaaa... Ya me lo contaste como mil veces.

—*Docza em bara*. ¿Entonces me lo cuentas tú?

—¡No! Otro, anda, tú conoces muchas historias.

—Sí, pero hoy no me acuerdo... A ver, anda, cuéntame tú el cuento de Ali Babá...

## Glosario

Palabra	Traducción literal	Traducción
osafah	Boca	la boca
cosi qicoh	Pequeña montaña	colina
coceqni (cosi ceqnu)	Pequeña mujer	niñita
bara meh	corazón yo	¿me lo juras?
seberqe	saber	Sabios
Auronev (aurok nevho)	Morir persona	Espíritu
faro anae kizna	Tú casa caminar	Sal de la casa
Ceeznee (ceez ceqnu)	Corteza mujer	Abuelita
vho kikaf ox	Tú ver negro	cierra los ojos
Barek (bara mek)	Corazón agua	tranquila
sebenhei dodaino xhuuxee (seberqe enhei dodax daino xhuuxee)	Saber cabeza qué nombre luna	Piensa en algo bonito
nvoque kikaf	Humo ver	No veo nada
Kxuve (kizna kgaxuve)	Casa frío	Cierra
Barenvho (bara enhei vho)	Corazón cabeza ojo	Prestar atención
Fivc ioroino (fivc iorosr daino)	Esto grande nombre	Esto es importante
kaomxee em devi aurok (kaomlan xhuuxee devi aurok)	Sol luna muerte camino muerte	Un día ya no voy a estar
beru kaomxee faro hecuz faro iosor egivaf (beru kaomlan xhuuxee faro hecu fuzd iosor egivo kikaf)	Largo sol luna tú tierra semilla tú grande escuchar ver	Falta mucho para que lo necesites, pero ya estás en edad de empezar a practicar.
em ioxo kafuh (em iorosr uhxo kikaf uhxo)	Yo grande uno ver uno	He estado muy sola sin ti.
evomlan emfarova uuxk (evoel kaomlan em faro vao uuxk)	Rato yo tú dos dormir	¿Puedo dormirme aquí contigo un rato?
Docza em bara	Eso yo corazón	Es que me gusta mucho

Tabla 4.1 Traducción de las frases utilizadas en el cuento

## Ejemplo de traducción

—*Nvoque kikaf, Ceeznee.* —Y abrió los ojos.

—*Kxuwen barenvo, fivc ioroino. Kaomxee em devi aurok...*— Su abuela le regresó la sonrisa.

—*Beru kaomxee faro hecuz faro iosor egivaf.* Ahora, lo que tienes que hacer...

—No veo nada, abuelita. —Y abrió los ojos.

—Ciérralos y presta atención. Un día ya no voy a estar...— Su abuela le regresó la sonrisa.

— Falta mucho para que lo necesites, pero ya estás en edad de empezar a practicar. Ahora, lo que tienes que hacer...

## 5 Conclusiones y futuros trabajos

Durante el proceso de la generación de palabras se definieron reglas pensando en un lenguaje en específico, el cual sería utilizado en una aldea, a la orilla de un río, y a no más de tres horas del mar. Se realizó una retroalimentación constante al programa de forma manual, ajustando los valores de evaluación para los sonidos de las letras y la fonotáctica, para encontrar valores que devolvieran palabras que pareciera provinieran de un mismo origen, o hayan sido creadas en un mismo contexto.

Se obtuvo un lenguaje que es fácil de pronunciar, pero difícil de hablar. A nivel fonético no hay problema, pero al trabajar la gramática, como se tomaron frases de diferentes idiomas, con diferente tipología, se complica un poco generar frases con facilidad. Es necesario tomar un tiempo para familiarizarse con la gramática, antes de poder hablarlo.

Este trabajo surge como una respuesta a la necesidad de crear un lenguaje artificial en un tiempo reducido. Acelerar un poco los procesos que los lingüistas utilizan para el diseño y definición de *artlangs*. Crear un lenguaje toma tiempo, sin embargo, en este trabajo de tesis se alcanzó este objetivo en un tiempo razonable considerando, a su vez, reglas fonotácticas del lenguaje. Y si bien existen generadores de palabras que ofrecen léxicos de forma aleatoria, el problema en estas soluciones es que no se tiene una cohesión entre palabras. Es necesario generar un gran número y tomar aquellas que se considere tengan una relación más estrecha unas entre otras. Creo que es aquí en donde está el primer enriquecimiento al proceso de creación de un lenguaje: los aspectos evolutivos de un lenguaje natural, la cruza y mutación de los caracteres. Los procesos evolutivos naturales del lenguaje de cierta manera ahora no se ven manipulados por una intervención externa, el lingüista, sino más bien creados en

función de probabilidades. Y reconociendo que ciertos procesos del lenguaje pueden incluso mejorarse, se pueden considerar otros criterios en la función de aptitud emulando el desuso de palabras, la integración de palabras de otros lenguajes, la creación de palabras a partir de fenómenos culturales, entre otras.

Este proyecto demuestra que los algoritmos evolutivos también pueden utilizarse para algo más creativo, algo que requiera cierta estructura y estética al mismo tiempo. No sustituye al creador, si no facilitar la parte más laboriosa. Incluso permite la iteración rápida, lo que permitiría explorar múltiples posibilidades que manualmente tomarían años.

Ahora, como trabajo futuro se tiene contemplada la generación de gramáticas, la cual, por cuestiones de tiempo, en este trabajo de tesis se realizó de manera “manual”. Sin embargo, esto es factible automatizarlo. Específicamente, los problemas se centraron en 1) delimitar el alcance de la gramática generada; 2) la recursividad del programa para poder reescribir las funciones de evaluación, que quizá se pudieran solventar a través de un *parser*; y 3) las posibles producciones con el mínimo de elementos (e.j. sustantivo, verbo, adverbio, adjetivo, etc) necesarios para que se entienda la idea en este nuevo lenguaje. Inicialmente, la idea era generar las reglas de producción para el nuevo lenguaje también utilizando programación genética, para después evaluar la capacidad de estas reglas para generar las oraciones que había elegido. Pero, trabajar con un lenguaje sin contexto, como no es lógico como las matemáticas, las posibilidades son infinitas al momento de definir reglas. Es un proceso de desarrollo más complejo porque no es un área que incluso se entiende en su totalidad. A la fecha Noam Chomsky, uno de los lingüistas más sobresalientes, reconoce que la gramática es indefinible, no hay definición única, correcta o incorrecta, para cualquiera de los lenguajes

existentes. Se tienen definidas reglas para las gramáticas actuales, pero como una idea de lo que podría pasar, a partir del análisis constante de la comunicación del ser humano; incluso así, estas reglas siempre pueden ser ignoradas y aun así podrían ser válidas.

Por eso, se decidió utilizar el algoritmo de GLR de forma manual para poder definir cuándo, y cuándo no, el algoritmo se tiene que detener.

En cuanto a trabajos futuros, es innegable el enriquecimiento que da el proceso evolutivo en la evolución de un lenguaje artificial. Por ello, la aplicación podría incluso llevarse más allá de los *artlangs*. Claros ejemplos son: ayudar en la definición de *auxlangs*; en la encriptación de datos, al generar de forma evolutiva las llaves públicas o privadas, o al generar cadenas de *hash* con características específicas; en la creación de lenguajes en clave para uso militar que permitan la comunicación con palabras cortas o incluso lenguajes completos que requieran de lingüistas enemigos para lograr descifrar el contenido de mensajes; en simulaciones de evolución lingüística, para llegar a los protolenguajes de ciertas lenguas, o a lo que podría llegar un lenguaje actuar con ciertos estímulos; en la composición musical, ya sea generando melodías o armonías que estén limitadas en notas, claves o un estilo de música; entre otras aplicaciones. Sin embargo, siempre con el objetivo de ser una herramienta que facilite la creatividad humana, no sustituyéndola.

## 6 Bibliografía

- [1] Adams, M. (2011). L4nGu4GE G4m35 in G4minG L4nGu4G35. *From Elvish to Klingon: exploring invented languages* (pp. 261-265). Oxford University Press
- [2] Aguilar Alconchel, M. Á. (2004). *Chomsky la gramática generativa*, en *Revista digital "Investigación y educación"*, vol. 3 (no. 7).
- [3] Anderson, J., & Staines, J. (2010). *A Genetic Algorithm for Grammars*.
- [4] Aycinena, M. A., Kochenderfer M. J. & Mulford D. C. (2003). *An evolutionary approach to natural language grammar induction*.
- [5] Back, T., & Fogel, D. B., & Michalewicz, Z. (1997). *Handbook of evolutionary computation*, IOP Publishing Ltd.
- [6] Bird S., & Loper E. (2004). *NLTK: The Natural Language Toolkit*. en *Proceedings of the ACL Interactive Poster and Demonstration Sessions*, pp. 214–217, Association for Computational Linguistics.
- [7] Charnaik, E. (1996). *Tree-bank Grammars*. Department of Computer Science, Brown University.
- [8] Chomsky, N. (1956). *Three models for the description of language*, en *IRE Transactions on Information Theory*, vol. 2 (no. 3), pp. 113–124.  
10.1109/TIT.1956.1056813

- [9] Chomsky, N. (2005). *Three Factors in Language Design*. vol. 36 (no. 1), pp. 1–22.
- [10] Emrys, S. (2006). *Ling 98/198: Constructing Languages DE-Cal*, Spring 2006, UC Berkeley By, a.k.a. Ilya Starikov <https://library.conlang.org/education/>
- [11] Emrys, S., & Fink A. & Peterson, D. (2009, Diciembre 28) *Conlanging 101*, <https://library.conlang.org/articles/journal.html>
- [12] Greenberg, J. H. (1974). *Language Typology: A Historical and Analytic Overview*.
- [13] Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press.
- [14] Instituto de las Personas Con Discapacidad (2017). *Diccionario de Lengua de Señas Mexicana de la Ciudad de México*.
- [15] International Phonetic association. (2020). *Full Kiel, Doulos, and DejaVu charts* [https://www.internationalphoneticassociation.org/IPAcharts/IPA\\_chart\\_orig/IPA\\_charts\\_E.html](https://www.internationalphoneticassociation.org/IPAcharts/IPA_chart_orig/IPA_charts_E.html)
- [16] Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty, and Daniel Zeman (2016). Universal Dependencies v1: A Multilingual Treebank Collection.

- [17] Junczys-Dowmunt, M. (2012). *A Genetic Programming Experiment in Natural Language Grammar Engineering*, en *Lecture Notes in Computer Science*. 10.1007/978-3-642-32790-2\_41.
- [18] Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection* (6th ed.). MIT Press.
- [19] Koza, John R. (1990). *Genetic programming: a paradigm for genetically breeding populations of computer programs to solve problems*. Stanford University.
- [20] Lang, B. (1974). *Deterministic techniques for efficient non-deterministic parsers*. en *International Colloquium on Automata, Languages, and Programming*. (pp. 255–269). Springer Berlin Heidelberg
- [21] Lucas, S. (1994). *Structuring chromosomes for context-free grammar evolution*. (pp. 130–135). 10.1109/ICEC.1994.350028.
- [22] LLorach, E. A. (1994). *Gramática de la lengua española*.
- [23] Meyer, Hans. (2016). *A global language or a world of languages*, en *Diffusion Fundamentals*, vol. 25, pp. 1–6. 10.62721/diffusion-fundamentals.25.906.
- [24] Nohejl, A. (2011). *Grammar-based genetic programming*.
- [25] Orkent, A. (2009). *In the Land of Invented Languages: Esperanto Rock Stars, Klingon Poets, Loglan Lovers, and The Mad Dreamers Who Tried to Build a Perfect Language*. Spiegel & Grau.

- [26] Prescher D., & Scha R., & Sima'an K., & Zollmann A. (2006). *What are treebank grammars?* en *BNAIC'06*.
- [27] Rosenfelder, M. (2010). *The Language Construction Kit*, Yonagu Books
- [28] Song, J. J. (2010), *The Oxford Handbook of Linguistic Typology*
- [29] Tomita M. (1984). *LR parsers for natural languages*. en *10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics*, pp. 354–357, Association for Computational Linguistics.
- [30] Firestrike (2025) *Tunic Game Wiki*. <https://tunic.wiki/books/faq/page/helpful-links>
- [31] van Oostendorp, M. (2001). *Constructed language and linguistic theory: on the notions possible, impossible, and actual language*, en *INTERFACE. Journal of Applied Linguistics*, vol. 15 (no. 2), pp. 203–221.
- [32] van Steenberg, Jan (2020). *Zonal Auxiliary Languages*. In I.K (Ed.) *The Role of Languages in Intercultural Communication / Rolo de lingvoj en interkultura komunikado / Rola języków w komunikacji międzykulturowej*. Cross-linguistic and Cross-cultural Studies 1. (pp.44). *Poznań*
- [33] Wyard, P. (1993). Context free grammar induction using genetic algorithms. *IEE Colloquium on Grammatical Inference: Theory, Applications and Alternatives*, IET, pp. P11–1.

# 7 Apéndice

## 7.1 Instrucciones para usar el código

Para correr el código en Python:

1. Es necesario tener instalado Python 3.8.
2. Poner el archivo **GenPro.py** y **requirements.txt** en una carpeta.
3. Posicionarse dentro de la carpeta
4. Abrir la ventana de comandos cmd y correr el comando:

```
pip install -r requirements.txt
```

5. Con ello, se puede correr el programa.
  - a. A través de la IDE de Python
    - i. Abrir la pestaña Run,
    - ii. Seleccionar Run Module
  - b. A través de ventana de comandos cmd
    - i. Correr el comando:  

```
Python GenPro.py
```

Para correr el código de Javascript:

1. Es necesario tener instalado NodeJS
2. Poner el archivo **index\_palabras.js** y **data.json** en una carpeta
3. Posicionarse dentro de la carpeta
4. Abrir la ventana de comandos cmd y correr el comando

5. Correr el comando:

```
npm install
```

6. Una vez finalizado, en la ventana de comandos correr el comando:

```
node index_palabras.js data.json
```