



Casa abierta al tiempo  
UNIVERSIDAD AUTÓNOMA METROPOLITANA

## ACTA DE EXAMEN DE GRADO

No. 00062  
Matrícula: 2143805478

MODELO FILOLÓGICO PARA  
LENGUAS FRANCESAS Y  
GERMÁNICAS UTILIZANDO  
COLORACIÓN DE GRÁFICAS  
SUAVES.

En la Ciudad de México, se presentaron a las 16:00 horas del día 2 del mes de junio del año 2017 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

DR. MIGUEL ANGEL GUTIERREZ ANDRADE  
DR. PEDRO LARA VELAZQUEZ  
DRA. HÉRICA SÁNCHEZ LARIOS



Bajo la Presidencia del primero y con carácter de Secretaria la última, se reunieron para proceder al Examen de Grado cuya denominación aparece al margen, para la obtención del grado de:

MAESTRO EN CIENCIAS (CIENCIAS Y TECNOLOGIAS DE LA INFORMACION)

DE: JORGE LUIS APARICIO REYES

y de acuerdo con el artículo 78 fracción III del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

Aprobar

Acto continuo, el presidente del jurado comunicó al interesado el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.

REVISÓ  
LIC. JULIO CESAR DE LARA ISASSI  
DIRECTOR DE SISTEMAS ESCOLARES

DIRECTOR DE LA DIVISIÓN DE CBI  
DR. JOSE GILBERTO CORDOBA HERRERA

PRESIDENTE  
DR. MIGUEL ANGEL GUTIERREZ ANDRADE

VOCAL  
DR. PEDRO LARA VELAZQUEZ

SECRETARIA  
DRA. HÉRICA SÁNCHEZ LARIOS

**Modelo filológico para lenguas  
francesas y germánicas utilizando  
coloración de gráficas suaves**

Tesis para obtener el grado de:  
**Maestro en Ciencias**  
(Ciencias y Tecnologías de la Información)

Presenta:  
**Lic. Jorge Luis Aparicio Reyes**

Asesores:

Dr. Pedro Lara Velázquez

Dr. Sergio Gerardo De Los Cobos Silva

Jurado calificador:

Presidente: Dr. Miguel Ángel Gutiérrez Andrade

UAM-I

Secretaria: Dra. Hérica Sánchez Larios

UNAM

Vocal: Dr. Pedro Lara Velázquez

UAM-I

Ciudad de México, Junio 2017

# **Modelo filológico para lenguas francesas y germánicas utilizando coloración de gráficas suaves**

Tesis para obtener el grado de:

**Maestro en Ciencias**

(Ciencias y Tecnologías de la Información)

Presenta:

**Lic. Jorge Luis Aparicio Reyes**

Asesores:

**Dr. Pedro Lara Velázquez**

**Dr. Sergio Gerardo De Los Cobos Silva**

---

## Resumen

En este trabajo se desarrolló un clasificador de lenguas germánicas y de Francia utilizando el modelo de coloración de gráficas suaves. Esta propuesta utiliza textos ASCII de cada uno de estos idiomas y genera comparaciones estadísticas de caracteres y obtiene una matriz de distancias entre lenguas.

Esta matriz conforma la información de entrada para el modelo de coloración de gráficas suaves, el cual genera tres características: dureza, solidez y resiliencia; este último valor nos indica cual es el número de colores (grupos o subfamilias) más adecuado para clasificar.

Los resultados del modelo se contrastaron con los obtenidos por los lingüistas durante décadas de trabajo multidisciplinario basados en estudios filológicos, hallazgos arqueológicos y análisis literarios, obteniendo resultados de la misma calidad, pero mediante un modelo de optimización matemática.

## **Agradecimientos**

Agradezco al CONACyT por la beca que me fue otorgada para poder realizar éste proyecto.

Al posgrado de CyTI porque confiaron en mis habilidades y me dieron la oportunidad de ingresar a la maestría.

A mis asesores Sergio Gerardo De Los Cobos Silva y a Pedro Lara Velázquez, el cual no sólo me asesoró y me dio buenos consejos, sino que también ha sido un buen amigo en los buenos y malos momentos.

A mis compañeros del posgrado por su solidaridad.

A mis amigos Erik Alarcón Zamora por sus consejos, su apoyo y todos los buenos consejos y Eugenia Méndez por su compañía tan cálida, las pláticas que hemos tenido y sus consejos invaluable.

A mi familia por todo su apoyo, en especial a mis tías Minerva Aparicio Reyes y Oliva Mejía Reyes quienes me apoyaron en los momentos más difíciles de mi vida.

Y sobre todo a mi mamá Ofelia Aparicio Reyes que me dio todo su amor incondicional y me enseñó todo de la vida, predicando siempre con sus buenas acciones. Todo se lo debo a ella...

# Contenido

|   |           |
|---|-----------|
| <b>1. INTRODUCCIÓN .....</b>  | <b>1</b>  |
| <b>2. OBJETIVOS .....</b>   | <b>3</b>  |
| <b>2.1 Objetivo General.....</b>  | <b>3</b>  |
| <b>2.2 Objetivos específicos .....</b>  | <b>3</b>  |
| <b>3. METODOLOGÍA .....</b>   | <b>4</b>  |
| <b>4. FAMILIA FRANCESA Y FAMILIA GERMÁNICA .....</b>                                | <b>5</b>  |
| <b>4.1 Familia Francesa.....</b>  | <b>5</b>  |
| 4.1.1 Historia del francés y las políticas en contra de las lenguas regionales..... | 7         |
| <b>4.2 Familia Germánica. ....</b>  | <b>35</b> |
| 4.2.1. Germánico Nórdico.....   | 40        |
| 4.2.2 Germánico Occidental .....  | 45        |
| <b>5. MARCO DE REFERENCIA.....</b>  | <b>56</b> |
| <b>5.1 Filología comparada y familias de lenguas.....</b>                           | <b>56</b> |
| <b>5.2 Modelo de coloración de gráficas suaves .....</b>                            | <b>57</b> |
| 5.2.1. Definición de coloración de gráficas suaves [Lara 2015].....                 | 57        |
| 5.2.2. Solidez y resiliencia de una coloración .....                                | 58        |
| 5.2.3. Distancia entre lenguas .....  | 59        |
| <b>5.3 Método de solución binario CGS [Lara 2015].....</b>                          | <b>59</b> |
| <b>6. ESTADO DEL ARTE .....</b>   | <b>61</b> |
| <b>6.1 Modelo Filológico para lenguas ibéricas.....</b>                             | <b>61</b> |
| <b>6.2 Reconocimiento de literatura inglesa .....</b>                               | <b>62</b> |
| <b>6.3 Un sistema Clasificador Utilizando Coloración de Graficas Suaves.....</b>    | <b>64</b> |
| <b>7. MODELO FILOLÓGICO PROPUESTO .....</b>   | <b>67</b> |
| <b>7.1 Algoritmo de Solución .....</b>  | <b>69</b> |

|   |           |
|---|-----------|
| <b>8. INSTANCIAS.....</b>                             | <b>73</b> |
| <b>8.1 Familia de lenguas francesas .....</b>         | <b>73</b> |
| <b>8.2 Familia germánica .....</b>                    | <b>75</b> |
| <b>9. RESULTADOS .....</b>                            | <b>78</b> |
| <b>9.1 Resultados para la familia francesa.....</b>   | <b>78</b> |
| <b>9.2 Resultados para la familia germánica .....</b> | <b>81</b> |
| <b>10. CONCLUSIONES.....</b>                          | <b>84</b> |
| <b>REFERENCIAS.....</b>                               | <b>86</b> |

## 1. Introducción

El trabajo de los lingüistas es tratar de explicar el origen de las lenguas actualmente habladas -entre estas las llamadas protolenguas-, de qué forma han ido evolucionando a través del tiempo, cómo se han ido influenciando entre sí y en qué manera surgen los idiomas basados en otros existentes. Algunos lingüistas creen que existió una protolengua que dio origen a todos los idiomas actuales, a la cual nombran como protosapiens. A esta teoría se le conoce como monogénesis. Por otro lado, la mayoría de los lingüistas creen que los lenguajes hablados surgieron en distintas partes del mundo a partir de distintas protolenguas, de este modo se adhieren a la teoría conocida como poligénesis. Quienes apoyan la monogénesis intentan recrear las distintas protolenguas del mundo y a partir de ellas tratan de reconstruir la protolengua protosapiens<sup>1</sup>, a pesar de que algunos expertos aseguran que no se pueden reconstruir lenguas con más de 5000 años de antigüedad y que el protosapiens podría tener una antigüedad de al menos 100,000 años.

Por lo anterior estos lingüistas deben realizar un arduo trabajo, que a menudo involucra diferentes disciplinas, así como diferentes procedimientos y fuentes. Una de ellas son los hallazgos arqueológicos, tales como construcciones, herramientas, vestimenta e incluso escritos hechos sobre diferentes materiales. Con estos hallazgos, los científicos pueden inferir acerca de las culturas y relaciones entre distintos pueblos, deducir los hábitos de sus habitantes, sus conocimientos y sus creencias. También es posible conocer las relaciones entre distintas civilizaciones como el comercio, intercambios culturales y conflictos. Otro tipo de fuente es la evidencia bibliográfica, que muestra claramente como han ido evolucionando las lenguas con el paso del tiempo y permite deducir los motivos de sus cambios.

La mayoría de las lenguas habladas actualmente en Europa, en algunos lugares de Asia y África, pertenecen a la familia protoindoeuropea. Esta a su vez se divide en varias subfamilias, entre las cuales se encuentran la germánica y la itálica. A los integrantes de esta familia también se les conoce como lenguas romances, algunas de las cuales son habladas actualmente en territorio francés, aunque no sean reconocidas oficialmente. Dentro de las

---

<sup>1</sup> William S-Y. Wang, James W. Minett, The invasion of language: emergence, change and death, Trends in Ecology & Evolution, Volume 20, Issue 5, May 2005, Pages 263-269

lenguas germánicas se tomarán en cuenta para los propósitos de este trabajo a aquellas que son más representativas en la actualidad. Esta rama a su vez se divide en tres grupos: la familia germánica oriental, germánica occidental y germánica nórdica o escandinava. Por otro lado, de las lenguas habladas en Francia solo se tomarán en cuenta aquellas que actualmente cuentan con el mayor número de hablantes. Estas lenguas se pueden clasificar básicamente en lenguas de *Oc* y lenguas de *Öil*.

A partir de la invención de la imprenta de tipos móviles<sup>2</sup> en la década de 1450 por el artesano alemán Johannes Gutenberg<sup>3</sup>, esta se volvió popular en Europa, debido a que la factura de los libros era más rápida y barata<sup>4</sup>. Poco después los impresores de la época empezaron a darse cuenta que había letras que se repetían con mayor frecuencia con respecto a las demás, con lo cual debían fabricar más tipos móviles. También notaron que, entre lenguas similares como el alemán y el neerlandés, la frecuencia en la que aparecía cada letra era muy similar en ambos idiomas y por el contrario entre lenguas que no se parecían como el español y el danés, estas frecuencias eran muy distintas. En este trabajo se busca aprovechar este hecho mediante la comparación estadística de caracteres ASCII para clasificar lenguas germánicas y de Francia por medio de un sistema que utiliza el modelo de coloración de gráficas suaves.

En este trabajo se hará un sistema clasificador de lenguas germánicas y de Francia, para esto se tomarán textos de los representantes de cada familia haciendo la comparación estadística de caracteres ASCII y finalmente utilizar un modelo de Coloración de Gráficas Suaves. El sistema arrojará las coloraciones óptimas que corresponden a cada número de colores, así como sus valores de dureza correspondientes. Con estos últimos se pueden calcular los valores de solidez y con estos los valores de resiliencia. Estos son los indicadores para saber cuáles son los números de colores (grupos o subfamilias) mejores para clasificar.

---

<sup>2</sup> Los tipos móviles son piezas hechas por lo general de metal, donde están las letras en relieve y negativo a imprimir.

<sup>3</sup> Su verdadero nombre fue: Johann Gensfleisch zur Laden zum Gutenberg.

<sup>4</sup> Aydon C. (2005) "Scientific Curiosity"

## 2. Objetivos

### 2.1 Objetivo General

- Crear un modelo filológico para la clasificación de lenguas germánicas y francesas utilizando Coloración de Gráficas Suaves.

### 2.2 Objetivos específicos

- Obtener un Estado del Arte sobre la Coloración de Gráficas Suaves.
- Investigar qué lenguas pueden ser factibles para coleccionar.
- Hacer una compilación de lenguas de la familia germánica.
- Hacer una compilación de lenguas habladas en Francia.
- Revisión bibliográfica y redacción sobre la historia y evolución de cada lengua utilizada.
- Diseñar y codificar un algoritmo para obtener matrices de distancias mediante los textos compilados.
- Diseñar y codificar un algoritmo para obtener las durezas y coloraciones óptimas para cada número de colores.
- Analizar los datos arrojados por el sistema para obtener las resiliencias para cada familia de lenguas.
- Reportar los resultados obtenidos mediante el sistema.
- Elaboración de las conclusiones y trabajo subsecuente

### 3. Metodología

Para la creación del Estado del Arte se revisaron tres trabajos relacionados donde se utilizaba la CGS, estos son: Modelo Filológico para lenguas ibéricas [Lara, 2015], Reconocimiento de literatura inglesa [Lara, 2015] y Un sistema Clasificador Utilizando Coloración de Graficas Suaves [Flores, 2016].

La obtención de los textos en cada uno de los idiomas es una parte fundamental de este trabajo debido a que esta tarea se tendría que realizar en un tiempo límite (máximo un trimestre) para poder avanzar en el resto del proyecto, por lo tanto, la investigación de factibilidad de los idiomas era primordial y los criterios principales fueron el número de hablantes de cada lengua, su disponibilidad en la web y su importancia histórica (como en el caso del gótico). Una vez hecho lo anterior se empezaron a coleccionar los textos para la familia germánica (alemán, danés, gótico, inglés, islandés, neerlandés, noruego *Bokmål*, noruego *Nynorsk* y sueco) y después se obtuvieron los correspondientes a los hablados en Francia (bretón, catalán, corso, francés, occitano, provenzal y vasco). A continuación, se comenzó a diseñar un algoritmo para la clasificación de las lenguas, este se divide en dos etapas: la primera analiza los textos de cada familia (germánica y francesa) y como salida se tiene una matriz de distancias en la cual se encuentran las penalizaciones para el modelo de CGS; en la segunda etapa, por medio de la matriz de distancias, se obtienen las durezas mínimas y sus respectivas coloraciones para cada número de colores (para la familia germánica se obtienen ocho valores óptimos de dureza mientras que son seis para las de Francia). Después se obtuvieron los valores de solidez y resiliencia a partir de los valores de dureza mediante el uso de una hoja de cálculo. Con los valores de resiliencia se puede decir cuales son los números más adecuados de grupos para clasificar.

Casi de manera paralela se realizó la revisión bibliográfica y la redacción sobre la historia y evolución de las lenguas incluidas en el trabajo, esto con la finalidad de conocer como fueron surgiendo estas, su evolución durante siglos y como han sido influenciadas unas con otras y conocer las clasificaciones hechas por los lingüistas, esto con el fin de saber que tan eficiente es este sistema ya que actualmente no existen trabajos con los cuales se pueda comparar.

## 4. Familia Francesa y Familia Germánica

### 4.1 Familia Francesa

Anterior al dominio romano en Europa Occidental, los pueblos celtas habitaron Galia actualmente gran parte de Francia. Las lenguas que estos hablaban eran inteligibles entre sí. En un principio estos pueblos vivieron en el este o centro de Europa y hacia el año 500 a.C. empezaron a migrar hacia el oeste, donde ocuparon Galia durante 200 años. Después empezaron una expansión hacia el oeste y sur de Europa, desplazando a los habitantes de Liguria (Italia) e incluso cruzaron los Pirineos hasta Iberia (España)<sup>5</sup>. En una pequeña zona entre Francia y España se situó la región de Aquitania (hoy en día conocida como Euskadi o País Vasco) en la cual se hablaba una lengua celta llamada aquitano, un tipo de euskera o vasco arcaico. Por otro lado, en la zona de *Massalia* (Marsella, Francia) se asentaron colonias griegas que hablaban y escribían en griego. Tanto los idiomas celtas como el griego desaparecieron de Galia tras la conquista romana liderada por Julio César en el año 52 a.C.

La expansión del Imperio Romano en gran parte de Europa trajo consigo cambios significativos, tales como la imposición a los pueblos conquistados de nuevas costumbres, religión, forma de vida, cultura y el latín como lengua oficial. Esto último con el fin de que los nuevos gobernadores pudieran dar a conocer sus leyes a los gobernados, por ejemplo, el cobro de impuestos y las sanciones a aquellas personas que no pagaran.

Algunos pueblos se resistieron a cambiar su lengua, como fue el caso de *Bretania*, donde se hablaba el bretón, y el País Vasco, donde se habla el euskera.

Tras el declive del Imperio Romano, y como parte de un acuerdo entre este último y diversas tribus germánicas, llegaron a Galia los pueblos francos y visigodos<sup>6</sup>. Los francos se asentaron al norte de Galia (hoy en día territorio del norte de Francia, Bélgica y Países Bajos) y al sur se asentaron los visigodos; En un principio ambos pueblos estaban separados por el río Loira y adoptaron el latín que era hablado por los galos. De esta manera, tanto los francos

---

<sup>5</sup>[http://orff.uc3m.es/bitstream/handle/10016/12343/celtas\\_lopez\\_2003.pdf;jsessionid=AE6FC3098B70A4FEB76E540ACD27434E?sequence=1](http://orff.uc3m.es/bitstream/handle/10016/12343/celtas_lopez_2003.pdf;jsessionid=AE6FC3098B70A4FEB76E540ACD27434E?sequence=1) [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history) Ball, Martin J. 2010. The celtic languages.

<sup>6</sup> <https://global.britannica.com/topic/Visigoth> <https://global.britannica.com/topic/Frank-people>

como los visigodos fueron adaptando el latín hasta convertirlo en dos lenguas distintas: a la lengua del norte se le llamó *Langue d'Oil* y la del sur *Langue d'Oc*<sup>7</sup>.

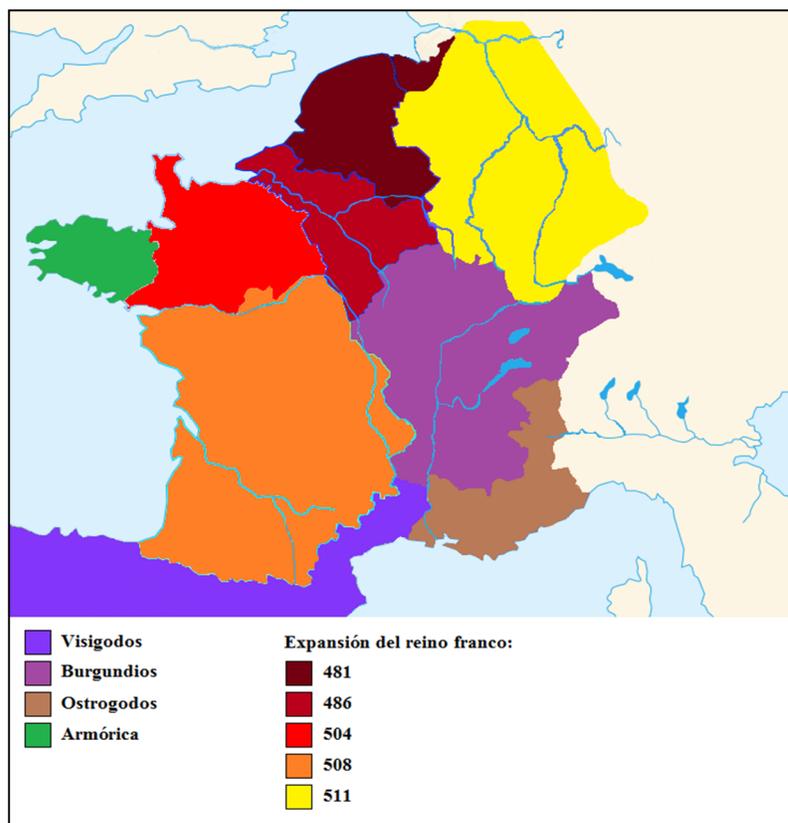


Figura 1. Distribución de las diferentes tribus germánicas y celtas en la Francia actual

Más tarde, en el año de 486 d.C., *Hlodowig* I (primer rey de los francos) venció a *Syagrius* (quien era rey romano de *Soissons*), con lo cual expande su territorio al sur y oeste de Galia y en consecuencia se da inicio a la monarquía de los merovingios, la cual unificaría a los pueblos francos<sup>8</sup>. En el año 508 *Hlodowig* designó a París como la capital de su reino y difundió la lengua de *Oil* por todo su reino. Con el paso del tiempo esta lengua se iría transformando hasta convertirse en el francés actual. *Hlodowig* hablaba el franconio o fránico antiguo (*vieux-francique*), una lengua de *Oil* que también sería la precursora del neerlandés.

<sup>7</sup> Las palabras *oil* y *oc* eran como se decía la palabra “si” en cada idioma.

<sup>8</sup> <https://global.britannica.com/biography/Clovis-I> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

Los expertos creen que el francés surgió entre los siglos VI y IX. Uno de los documentos más antiguos de este idioma es el de los *Juramentos de Estrasburgo* del año 842<sup>9</sup> d.C. Poco después, este idioma fue ampliamente utilizado por los clérigos (hombres relacionados con la iglesia) quienes empiezan a escribir las primeras obras literarias de esta lengua, una de estas es *La Secuencia de Santa Eulalia* (contenida en una colección de discursos, escrita en el año 880 d.C. por San Gregorio Nacianceno) que fue la primera obra escrita en francés antiguo. También en esta época se escribió *El Cantar de Roldán*, el cual es un poema épico de finales del siglo XI y que es atribuido a un monje normando. *El Cantar de Roldán* es considerado tan importante como *El Cantar del Mio Cid* (español), *El Cantar de los Nibelungos* (*Nibelungenlied*) germano y *Beowulf* (inglés).

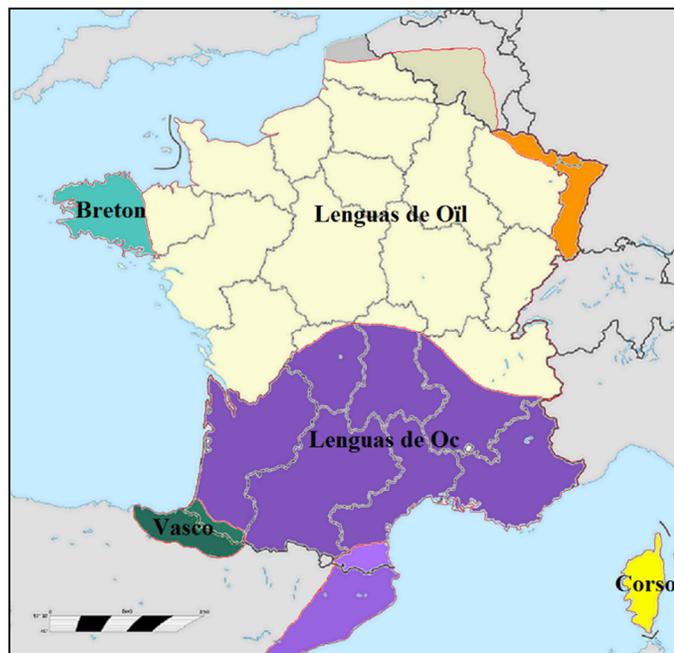


Figura 2. Distribución geográfica de las lenguas habladas actualmente en Francia

#### 4.1.1 Historia del francés y las políticas en contra de las lenguas regionales

Los orígenes de los pueblos franceses se remontan cerca de un millón de años en el pasado, según evidencia arqueológica. No obstante, los primeros pueblos de los que se tiene registro fueron los celtas, pueblos guerreros que invadieron y se asentaron en Europa

<sup>9</sup> [http://w3.restena.lu/cul/BABEL/T\\_SERMENTS.html](http://w3.restena.lu/cul/BABEL/T_SERMENTS.html) D'Arco Silvio Avalle, Luciana Borghi Cedrini. 1966. Alle origini della letteratura francese: i giuramenti di Strasburgo e la sequenza di Santa Eulalia

Occidental a partir del siglo V a.C. y hasta la llegada del Imperio romano alrededor del siglo I a. C. Estas tribus probablemente llegaron desde el río Danubio abarcando hasta partes como Italia y la península Ibérica. Al llegar a territorio de la actual Francia, aproximadamente en el siglo III a. C., se encontraron con otros pueblos que no opusieron resistencia a la invasión; como ocurre en muchos casos, los pobladores originales adoptaron las costumbres de los invasores, en este caso las costumbres de los celtas y se mezclaron con estos. Los pobladores se agruparon en villas compuestas en su mayor parte por individuos que tenían el papel de campesinos y guerreros a la vez; estos formaban tribus cuya sociedad se dividía en distintas clases. La población de galos (llamados así por los griegos que se asentaron en Marsella en el siglo VII a. C.) era de unas 10 millones de personas, que se dividieron en más de cien tribus, entre las cuales se encuentran los *aquitani*, *armoricae*, *arveni*, *britons*, *burgundi*, *cataluni*, *elyseens*, *franci*, *galates*, *helvetii*, *parisii*, *remi*, *suevi*, *vascones*, *veneti*, etcétera.

Debido a la gran cantidad de pueblos celtas, era lógico que su lengua se fragmentara en varios dialectos. La comunicación entre pueblos vecinos era sencilla debido a que estaban en contacto constante, mientras que para los pueblos más distantes entre sí no lo era, ya que sus lenguas eran muy distintas.<sup>10</sup> La mayor parte de la comunicación de los galos era oral y rara vez lo hacían de manera escrita, adoptando la escritura de los pueblos asentados en Marsella<sup>11</sup>.

El territorio de Galia era muy extenso ya que, como se mencionaba anteriormente, cubría el territorio de Francia, Suiza, Bélgica y la parte occidental de Alemania.

---

<sup>10</sup> Ninguna de las lenguas celtas habladas en la actualidad como el bretón, galés, irlandés, escocés, entre otros, surgieron de las lenguas habladas por los galos.

<sup>11</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)



Figura 3. Territorios galos (Galia<sup>12</sup>).

Con la conquista de Galia, los romanos impusieron varias de sus costumbres, religión y de alguna manera el latín porque ignoraban a la gente que hablaba galo, con lo que hablar latín se volvió indispensable para los nativos, pues esto les permitía trabajar para el Imperio romano y pretender tener una mejor calidad de vida.<sup>13</sup>

Además el Imperio, como pago por los servicios prestados, les otorgaba tierras a los romanos. Los beneficiados empezaron a construir caminos entre sus pueblos para poder llegar a cualquier parte del Imperio, con lo que cualquier romano podía tener acceso a lugares remotos de Galia. Esto tuvo como consecuencia la difusión del latín por todo el territorio galo, dando como resultado que todos los galos fueran bilingües, aunque poco a poco las lenguas celtas quedaron en desuso por motivos prácticos hasta el siglo V<sup>14</sup>, dejando solo unas 150 palabras que retomó el latín, estas palabras correspondían principalmente a la agricultura,

<sup>12</sup> [https://upload.wikimedia.org/wikipedia/commons/2/2b/Peuples\\_gaulois.jpg](https://upload.wikimedia.org/wikipedia/commons/2/2b/Peuples_gaulois.jpg)

<sup>13</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history) 10/Enero/2016 18:25

<sup>14</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

objetos usados en la vida diaria, a la fauna y a la flora, por ejemplo, *alauda* (alondra), *cervesia* (cerveza), *druid* (druida), *bard* (bardo, poeta), etc.

A pesar de que se extinguieron una gran variedad de lenguas en el territorio galo, el latín se empezó a dividir en dos dialectos, uno era el llamado latín clásico que era hablado por la gente con una mayor posición social y enseñado a estos para que lo hablaran correctamente. El otro tipo de latín era hablado por el resto de la población en su vida cotidiana, de modo que al ser hablado por la mayoría pasó a ser la variante utilizada para los actos religiosos y para la redacción de distintos documentos, con lo que el latín vulgar sustituyó al latín clásico.

Por otro lado, a partir del año 375 (fecha en que los ostrogodos fueron vencidos por los hunos) diversas tribus germánicas empezaron a invadir el occidente de Europa: los visigodos invadieron gran parte de España junto con el sur de Francia, los suevos tomaron el norte de España y Portugal, los francos se asentaron al norte de Francia, los burgundios habitaron el oriente de Francia y Suiza, los alemani se situaron en Alemania, los anglos y los sajones migraron hacia el Reino Unido, los ostrogodos se establecieron en Italia y Austria, finalmente los vándalos colonizaron las islas de Cerdeña y Corcega, así como el norte de África.

Hacia finales del siglo V había asentamientos de tribus germánicas bajo el tratado de *foederati*, en el cual se establecía que podían permanecer en territorio romano, además de recibir dinero y alimento, sin embargo, no eran considerados ciudadanos romanos ni colonias romanas, a cambio de los beneficios de los cuales gozaban, tenían que participar en batallas junto con los romanos siempre que estos lo solicitaran<sup>15</sup>.

Tiempo después, cuando los ingresos por impuestos fueron decayendo, el Imperio permitió a los germanos crear villas y de alguna manera se les empezó a considerar como ciudadanos, esto con el propósito de recaudar más impuestos. Se crearon varias villas autosuficientes, es decir, no necesitaban del apoyo del Imperio para la creación de la infraestructura que necesitaban y para el mantenimiento de la misma. Otro fenómeno fue la fragmentación de la mayoría de las tribus. La mayoría de los germanos estaban hartos del

---

<sup>15</sup> [http://penelope.uchicago.edu/Thayer/E/Roman/Texts/secondary/SMIGRA\\*/Foederatae\\_Civitates.html](http://penelope.uchicago.edu/Thayer/E/Roman/Texts/secondary/SMIGRA*/Foederatae_Civitates.html)  
<http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.04.0062:entry=foederatae-civitates-harpers>

Imperio debido a que les seguía cobrando impuestos a pesar de tener villas autosuficientes, por lo que se empezaron a rebelar, como en el caso de los visigodos que derrotaron a los romanos en la batalla de Adrianópolis y los francos en la batalla de *Soissons* en el año 486 liderados por el rey Clodoveo I<sup>16</sup> contra las tropas de Afranio Siagrio, con lo que se fundó el reino franco extendiendo aún más su territorio.

Los distintos pueblos germanos asentados por Europa Occidental tuvieron que aprender latín, que eran la lengua de los romanos, con el paso del tiempo surgieron varios dialectos del latín (latín vulgar) que fueron evolucionando hasta convertirse en las lenguas romances de la actualidad. [Charles, H.G. 1907]

*El satiricón* es una novela en verso y prosa, escrita por el político romano Gaius Petronius Arbitr. Su importancia reside en que muestra la forma de hablar de la gente del siglo I, aunque el autor narra hechos de su obra en latín clásico, sus personajes utilizan el llamado latín vulgar que utilizaba la mayoría de la gente.

Otra obra importante que muestra el surgimiento del latín vulgar, es el *Appendix Probi*, en donde se puede observar una compilación de errores frecuentes en el uso del latín. Se decía que esta compilación fue realizada por Marco Valerio Probo, pero actualmente se descarta su autoría. En esta obra se habla de los usos de los sustantivos, las reglas de ortografía y una tabla de diferencias o errores comunes del latín del siglo III, en el cual se sigue el modelo A non B. Algunos ejemplos son *Passim non Passi*, *Anfora non Ampora*, *NVMQVAM NON NUMQVA* (*Numquam non Numqua*). Este tipo de errores eran cometidos y aceptados por la mayoría de los ciudadanos romanos quienes fueron modificando paulatinamente su vocabulario, mientras tanto se fueron creando dialectos del latín en diversas regiones del Imperio Romano, hasta el punto en que las diferencias en los vocabularios fueron tan notorias que a estos dialectos se les consideró lenguas distintas (actualmente son las lenguas de la familia romance). Un ejemplo de la evolución de las palabras se puede apreciar en la palabra en latín “Numquam” que evolucionó a la palabra en latín vulgar “Numqua” que a su vez evolucionó a la palabra en castellano “nunca”.

---

<sup>16</sup> El nombre Clodoveo es una castellanización de *Hluot-wig*, *Hluot* reconocido y *wig* combate en bajo francico, *Hluot-wig* se convertiría en *Chlodwig* y *Hludwig* que a su vez darían origen a diversos nombres como *Ludwig* en alemán y Luis, por lo que la mayoría de los reyes franceses serían nombrados así.

### ***La dinastía de los Merovingios***

El plan de Clodoveo era expandir su reino invadiendo y conquistando otros territorios para después repartirlos entre su descendencia. De esta manera en 486 ganó la batalla de *Soissons* derrotando al último *Magister Militum*, el más alto jefe militar romano de esta región -que comprendía desde el río Mosa hasta Loira-, con lo cual el reino de Clodoveo se extendió desde el río Loira hasta el Rín y convirtió a dicha ciudad en la primera capital del reino de los francos. Clodoveo se casó con Cleotilde, hija de Chilperico II de Burgundia, quien por varios años le suplicó a Clodoveo que se volviera católico. En 496 este le juró a su esposa hacerlo en caso de ganar la batalla de *Tolbiac* contra los alamanes, y tras ganar la batalla, él, su hermana Lanthilde y 3 mil de sus hombres se bautizaron en Reims, lo que haría a Francia católica desde sus inicios. Además, Clodoveo I fue el primer rey de Francia en hablar el francico antiguo (*vieux-francique*), una lengua germánica que sería hablada por toda la dinastía de los Merovingios y los Carolingios, y desplazada de la realeza por el francés antiguo con el ascenso de Hugo Capeto en 987, el primer rey de la dinastía de los Capeto.

A pesar de tener el dominio de gran parte de Galia, únicamente el 5% de los habitantes eran francos, mientras que la mayoría de los pobladores hablaban una lengua romance llamada galo-romano, misma que contaba con varios dialectos regionales; algunos galos hablaban otras lenguas como el alsaciano, bretón, flemish y vasco. Sin embargo, los francos influenciaron el galo-romano con el uso del acento nórdico y un sistema fonético que permitía tener una pronunciación más marcada de las vocales, con lo que inició la germanización del galo-romano. De esta manera, durante el periodo de Carlomagno, la aristocracia franca fue bilingüe durante siglos antes de adoptar dicha lengua. El francico antiguo fue quedando gradualmente en desuso entre la realeza, sin embargo, continuaría siendo usada por los descendientes de los francos, hasta convertirse en neerlandés antiguo. Dos tercios de la población gala hablaba algún dialecto galo-romano germanizado, excepto la aristocracia, los pobladores de Bretaña, el País Vasco y noreste de Francia<sup>17</sup>. Los lingüistas afirman que en el francés actual se mantienen unos centenares de palabras tomadas del francico antiguo<sup>18</sup>. Además de haber tomado prestadas algunas palabras de este, el francés también tuvo otro

---

<sup>17</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

<sup>18</sup> La lingüista Henriette Walter afirma que son exactamente 544 palabras. Walter, H. (1997). *L'aventure des mots français venus d'ailleurs*.

tipo de influencias del francico, como la fonética, morfología y sintáxis. Algunos ejemplos de sintaxis germana se encuentran en las palabras *Rougemont* (monte rojo), *Clairvaux* (valle claro) o *Trouville* (ciudad agujero); en estos el adjetivo antecede al sustantivo, como ocurre en las lenguas germánicas. En el caso de la morfología -las reglas lingüísticas de flexión, composición y derivación de las palabras- la letra “h” se le atribuye al francico, debido a que en el latín esta letra se dejó de pronunciar desde la época de Cicerón, pero con la influencia germana se incluyeron palabras como *huit* (ocho), *huis* (puerta), *hermine* (ermitaño), *huître* (ostra), etc. Incluso el nombre de Francia<sup>19</sup> proviene de los francos y actualmente, en lenguas germánicas el nombre de Francia quiere decir literalmente “Reino o tierra de los francos”<sup>20</sup>.

Lo anterior muestra que el francés evolucionó de manera distinta a las demás lenguas romances e incluso puede ser considerada la lengua menos romance de este grupo. El galorromano continuó evolucionando, pero como no había ninguna organización que lo regulara, se dividió en distintas variantes locales. Lo que hoy en día se conoce como francés antiguo (siglos IX al XIV) era entonces un grupo de variantes orales que no se estandarizaron o codificaron. Los grupos de lenguas francesas se empezaron a dividir principalmente en tres grupos: las lenguas de *Oïl* al norte de Francia, las lenguas de *Oc* al sur y el grupo franco-provenzal al oriente. Cabe señalar la presencia de otros grupos, como el de lenguas de *Sì* en Italia y las islas de Cerdeña y Corcega que actualmente son territorio francés. Las lenguas de *Oïl* al norte del río Loira darían origen al francés, como en el caso del franciano (o *francien*, una lengua romance-germánica distinta al francico) hablada en la región de la *Île-de-France* (con París en el centro) entre los siglos IX y X. Por otro lado, se encontraban las lenguas de *Oc* al sur del río Loira, que aunque no estuvieron bajo el reino franco, fueron parte del reino de los visigodos, otra tribu germana. Estas lenguas eran más parecidas a otras lenguas romances debido a que no tuvieron mucho contacto con las lenguas de *Oïl*. Finalmente, los hablantes de lenguas provenzales fueron gobernados por los burgundios, también germanos. Por lo tanto, se puede observar que existía un gran número de lenguas en lo que hoy es

---

<sup>19</sup> El nombre de Francia proviene de la palabra en fránico *Frank* que quiere decir “libre”, de esta se desprenden las palabras *free* en inglés, *frei* en alemán, que tienen el mismo significado e incluso la palabra castellana “franco” tiene entre sus sinónimos a la palabra “libre”. La palabra “Francia” proviene del latín clásico y sobrevive en el castellano, gallego e italiano, sin embargo, en latín cambió a *Francorum*.

<sup>20</sup> Francia en alemán se escribe *Frankreich*, en danés *Frankrig*, en islandés *Frakkland*, en neerlandés *Frankrijk*, en noruego y sueco *Frankrike*.

Francia, así que las personas utilizaban el latín de la iglesia para poderse comunicar con habitantes de distintas regiones.

Hay que señalar que las clasificaciones de lenguas de *Oïl*, *Oc* y *Si* se deben al poeta Dante Alighieri, nacido en la Toscana durante el siglo XIII. En una de sus obras llamada *De vulgari eloquentia*, entre otras cosas, clasifica a varias lenguas romances según la manera de afirmar de cada región. La palabra *Oc* proviene de la palabra latina *Hoc* que literalmente quiere decir “eso”, y la palabra *Oïl* también proviene de la palabra *Hoc*, pero se fue modificando para darle más énfasis a la afirmación, mientras que la palabra *sì* en italiano significa “sí”.

### ***Cronología del francés antiguo.***

En 782, Carlomagno mando a llamar a Alcuino de Inglaterra para que formara parte de su corte, enseñara latín y para latinizar al francés.

El primer documento en francés fue llamado *Serments de Strasburg* o *Los juramentos de Estrasburgo*, redactado el 14 de febrero de 842 por los nietos de Carlomagno, Carlos el Calvo y Luis el Germánico, y transcrito por otro de sus nietos -Nitardo servidor de Carlos- al francés antiguo para que fuera entendido por la mayoría de la gente. Este documento fue redactado originalmente en latín literario por parte de Luis para que fuera entendido por su hermano Carlos y este lo redactó en *teudisca lingua* (lengua germana). En este documento se pactó una alianza en contra de Lotario I, rey de Francia media, emperador de occidente, rey de Italia y hermano mayor de ambos, debido a inconformidades con posesiones que les heredó su padre Luis el Piadoso. Luis el Germánico era mayordomo de Baviera y Sajonia, mientras Carlos el Calvo era mayordomo de Borgoña y Normandía. Tras la guerra civil carolingia, los tres hermanos firmaron el *tratado de Verdún*, con lo que oficialmente daban por terminado el conflicto, además de quedar por escrita la repartición de las tierras heredadas de manera más equitativa. De este modo, Carlos el Calvo recibiría Francia occidental (actualmente Francia), Lotario I se quedaría con Francia media (hoy en día Países Bajos, Suiza, Luxemburgo, Bélgica, oeste de Alemania y norte de Italia) y Luis el Germánico obtendría Francia oriental (ducados de Alemannia, Baviera, Sajonia y Turingia, territorios de la Alemania actual).

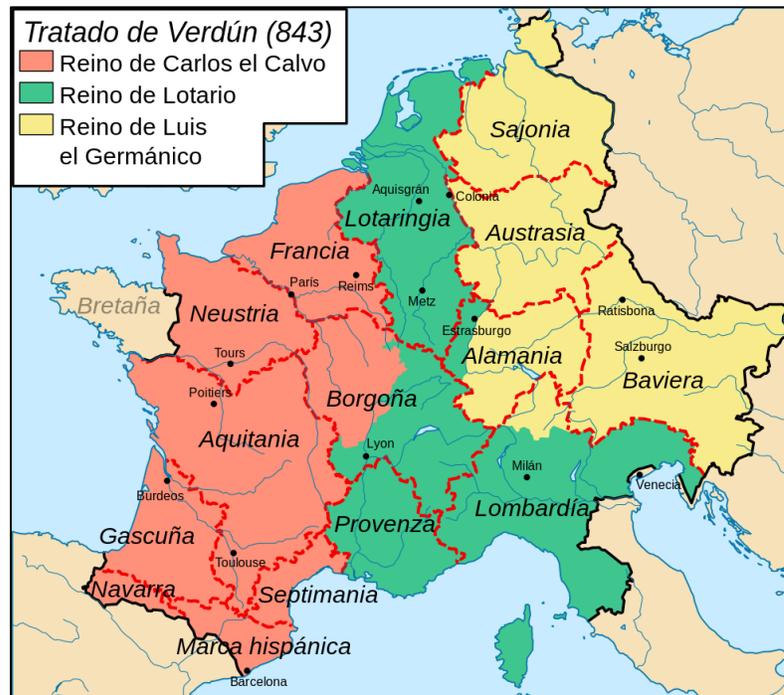


Figura 4. Repartición de territorios de acuerdo al tratado de Verdún.<sup>21</sup>

Durante el siglo siguiente, Hugo Capeto era una de las personas más influyentes de Francia (incluso más que el rey) debido a que poseía muchos territorios, principalmente en los alrededores de París que ya era la más importante ciudad del reino. Desde el abuelo de este, los Capeto ya pretendían el trono, no obstante, se mantenían con un perfil bajo, esperando la oportunidad de obtenerlo legalmente, es decir, siendo elegidos. Finalmente, la espera terminó para Hugo Capeto. Con la muerte de Luis V, El Holgazán en mayo de 987, último rey Carolingio, quien no dejó descendencia que continuara su reinado, con lo que el trono quedó desierto. El único miembro de esta familia que podía gobernar (su tío Carlos) no contaba con la simpatía de clérigos y señores feudales de Francia, mientras que Hugo ya se había aliado secretamente con Adalberón, arzobispo de Reims, y su pupilo Gerberto de Aurillac, pues estos habían tenido rencillas con la familia Carolingia. Juntos convencieron a los otros señores feudales de Francia para elegir a Hugo como su nuevo soberano. Solo unos días después de la muerte de Luis V, se reúnen los señores feudales para deliberar quién debía

<sup>21</sup> [https://es.wikipedia.org/wiki/Archivo:Verdun\\_Treaty\\_843-es.svg](https://es.wikipedia.org/wiki/Archivo:Verdun_Treaty_843-es.svg)

subir al trono. Finalmente, Hugo Capeto es elegido el 1 de junio de 987 y coronado el 3 de julio del mismo año.<sup>22</sup>

Hugo Capeto fue el primer rey de “Francia” que hablaba una lengua distinta al francico antiguo, pues hablaba una de las lenguas de *Oïl* de la región de la *Île-de-France*, que se podría decir que era francés antiguo.<sup>23</sup> La dinastía de los Capeto reinó con mayor poder que la anterior, extendiendo su reino hacia el sur.

En 1119 (a once años de su coronación), el rey Luis VI le escribió una carta al papa Calisto II proclamándose como “Rey de Francia” y no “El rey de los francos” como lo harían sus antecesores. Esta carta es la primera referencia escrita de Francia como reino.

Poco antes del siglo XII, en Francia las personas se comunicaban casi de manera exclusiva con personas de su misma región, haciendo la comunicación “vertical”, es decir, las clases sociales marcaban la forma de hablar de las personas, sin embargo, a inicios del siglo XII la comunicación se volvió “horizontal”, es decir, que la diferencia entre las lenguas se volvió regional, debido a las peregrinaciones, cruzadas, etc.

El francés antiguo fue una lengua vehicular (*lingua franca*) llevada a Italia, España, Chipre, Siria y Jerusalén durante las Cruzadas (entre los siglos XI y XIII). Posteriormente fue utilizado en forma escrita para la administración real junto con el latín, que seguía manteniendo cierto prestigio en Francia. El francés, al ser cada vez más utilizado en todo el reino y no solo por la aristocracia en la región de la *Île-de-France*, perdió parte de dicho prestigio, con lo cual ya no sería considerado como la lengua de la cultura<sup>24</sup> y por lo tanto el latín permanecería como la lengua utilizada para la educación, la justicia y las cancillerías excepto entre Francia e Inglaterra puesto que el francés era la lengua oficial del Reino Unido.

---

<sup>22</sup> Asimov, Isaac. 1972. La formación de Francia.

<sup>23</sup> El francés antiguo era un conjunto de lenguas de *Oïl* que fueron la base para crear el francés hablado en el siglo IX al XIV.

<sup>24</sup> Los intelectuales de la época, al decidir que el francés ya no sería considerado una lengua culta, empezaron a usar un francés más latinizado.

### ***Época del renacimiento y el francés medio.***

Los siglos XIV y XV fueron muy duros para Francia por diferentes motivos, como las hambrunas, las plaga, la peste negra, guerras civiles y la “Guerra de los Cien Años” contra los ingleses; no obstante, en esta época se inició el desarrollo del francés medio.

Durante el periodo de la Guerra de los Cien Años, el francés experimentó muchos cambios en su pronunciación, tales como la simplificación, eliminando los diptongos y triptongos, puesto que la gente culta trató de preservar la lengua escrita, aunque ya no correspondiera con la hablada. Esto se puede apreciar en palabras del francés de la actualidad como “oiseau” (pajaro) pronunciado simplemente como “wazo”, “peau” (piel) pronunciado como “po”, “saoul” (borracho) pronunciado como “su”, etc. No sólo cambió la pronunciación de varias palabras, sino que también se vieron afectadas algunas reglas gramaticales simplificando al francés hablado, mientras que el escrito se mantuvo inalterado con respecto a su ortografía, por lo que se puede decir que es una especie de latín afrancesado. Otro fenómeno que se dio en esta época fue la introducción de préstamos lingüísticos del latín al vocabulario francés por parte de la gente culta. Aunque muchas de estas palabras tuvieron un tiempo de vida muy breve (como fue el caso de *intellectif*, *médianable*, *suppécliter*, etc.), otras permanecen hasta el día de hoy (como *déduction*, *altercation*, *incarcération*, *prémisse*, etc). Este mismo fenómeno ocurriría dos siglos más tarde con el inglés isabelino, que tomo muchos préstamos del latín y griego, además se observa un intento por crear nuevas palabras con raíces germanas para contrarrestar la influencia romance del francés. Otro fenómeno más fue el de las palabras dobles, que son palabras similares debido a que tienen las mismas raíces etimológicas con la diferencia que una de ellas fue evolucionando naturalmente con el francés y la otra fue tomada directamente del latín. Algunos ejemplos son las palabras *régalien* y *regalen* (de la realeza) o *gourde* y *couge* (calabaza).

La última etapa del francés medio (siglo XVI) se caracterizó por tomar prestamos lingüísticos del italiano, pues Italia pasaba por un apogeo en todos los aspectos y por consiguiente el italiano era considerada como una lengua con gran prestigio. Muchos italianos aprovecharon esta situación para hacer carreras en la corte del rey, o posicionarse como intelectuales, artistas y científicos. En total se tomaron alrededor de 8000 palabras del italiano, desde palabras de uso cotidiano hasta palabras del argot de distintas profesiones,

áreas del conocimiento y de las artes. De estas palabras se continúan usando el 10%, por ejemplo, *arcade* (arcada, palabra introducida en el año 1562 que significa conjunto o serie de arcos), *balcon* (balcón, introducida en el año 1570), *carrousel* (carrusel, introducida en el año 1596), *duo* (dúo, introducida en 1547), *escarpe* (escarpa, declive, plano inclinado, introducida en 1549)<sup>25</sup>, etc.

Durante el siglo XVI el francés, frecuentemente llamado lengua común, ganó cada vez más importancia en Francia. En el año de 1510 el rey Luis XII decretó que los procedimientos judiciales se escribirían en “lengua común” en detrimento del latín y en 1539 el rey Francisco I firmó la ordenanza de *Villers-Cotterêts* redactada por Guillaume Poyet. Este es el texto legislativo más antiguo en vigencia en Francia. Este documento, entre otras cosas, obliga a que todos los documentos oficiales -ya sean administrativos o judiciales- se deben redactar en la lengua común, dejando en desuso al latín y las lenguas vernáculas y por lo tanto el francés sería la lengua oficial de Francia. Lo anterior logró oficializar el uso del francés como parte del movimiento de centralización lingüística que se había iniciado desde el siglo XIV. También los tipógrafos impulsaron la difusión del francés, pues cada vez más gente prefería leer en este idioma que en latín. No sólo se limitaron a difundir la lengua, sino que también son los causantes de su complicada ortografía, debido a que eran pagados según la extensión de las palabras a imprimir en los libros; además se les atribuye el uso de distintos tipos de acentos, el apóstrofo y la cedilla.

### ***El francés clásico***

El francés clásico es el antecesor inmediato del francés moderno, donde se empieza a estandarizar debido a la falta de reglas para el uso del francés. Así es como en 1647 Claude Favre de Vaugelas, gramático y miembro de la *Académie française*, publicó su obra *Remarques sur la langue française, utiles à ceux qui veulent bien parler et bien écrire*, que era una colección de observaciones de gramática y lexicográficas basadas en la manera de hablar de la corte. Vaugelas insistió en el prefacio que su libro era un recordatorio, pero no una imposición a la forma de hablar. También, durante este periodo se dejó a un lado el chauvinismo o el nacionalismo exagerado con respecto a la lengua, por lo cual se incluyeron

---

<sup>25</sup> <http://www.cnrtl.fr/etymologie>

muchos prestamos lingüísticos, por ejemplo, del italiano las palabras: *attitude, calque, coloris, filigrane, fresque, groupe, miniature, morbide, morbidesse, reflet, svelte, basson, opéra, partition, récitatif, ritournelle, sonate, ténor, bandit, café, céleri, cortège, désinvolve, escompter, lésine, mousseline, pantalon, polichinelle, réussite, salon, volte-face*, etc.; del inglés: *yacht, paquebot, rhum, flanelle, y moire*; del español (España): *matamore y romance*; de las colonias españolas: *chocolat y tabac*; del alemán: *blocus, bivouac y sabre*.<sup>26</sup>

Esta etapa surgió en el *Grand Siècle* (el gran siglo) el cual fue una época caracterizada por una prosperidad económica, estabilidad social y un alza en el prestigio de la política, literatura y las artes, dejando atrás los años difíciles. Todo gracias al rey Enrique IV (apodado “el grande” o “el buen rey”).

El francés moderno empezó a surgir en el periodo del *Grand Siècle* (el gran siglo) el cual fue una época caracterizada por una prosperidad económica, estabilidad social y un alza en el prestigio de la política, literatura y las artes, dejando atrás los años difíciles. Todo gracias al rey Enrique IV (apodado “el grande” o “el buen rey”).

Durante el periodo del *Grand Siècle* la autoridad se centró cada vez más en los reyes que estuvieron en turno, empezando con uno de los reyes más queridos -sino el más querido- de la historia de Francia: Enrique IV, quien se preocupó por su pueblo, fomentando la educación, haciendo varias obras públicas, promoviendo la agricultura, etc. Esto no solo le favoreció personalmente, sino también a su descendencia. Durante el reinado de Luis XIII, apodado “el justo”, surgió un personaje muy influyente en la historia de Francia: el cardenal Richelieu, quien además de crear un estado centralizado<sup>27</sup>, también creó la *Académie française* en 1635, la cual es hasta la fecha la sociedad principal literaria francesa y que promueve el uso del francés y su literatura. Durante el siglo XVII, el francés alcanzó la perfección, según los gramáticos de la época. Ellos fueron los encargados de diseñar esta lengua a su gusto, seleccionando vocablos elegantes, además de eliminar muchas otras palabras consideradas impuras, como fue el caso de los arcaísmos, italianismos (tomados un

---

<sup>26</sup> Rickard P. (2003). “A history of the french language”

<sup>27</sup> Antes cada territorio tenía sus propias leyes y los señores feudales a veces tenían incluso mayor influencia que el rey y disponían de sus propios ejércitos, con lo que muchos de ellos frecuentemente trataban de conspirar en contra del rey en turno. Sin embargo, al crearse un estado centralizado o unitario Francia solo tenía una constitución y las mismas leyes aplicaban para toda la nación, restaurando la autoridad real.

siglo antes), palabras del *patois*<sup>28</sup> y términos técnicos de distintas profesiones, además se dejó de pronunciar la letra “s” en las palabras en plural. Una vez hecho lo anterior, la *Académie française* publicó su primer diccionario en 1694. A pesar del perfeccionamiento del francés, solo el 1% de los habitantes lo hablaba y el resto se comunicaba a través de alguna lengua local, aunque poco a poco esta gente empezaría a hablar francés sin la elegancia esperada. Durante sus viajes a través de Francia, el rey Luis XVI tenía que dar sus discursos en las diferentes lenguas de los territorios que visitó. Los territorios recientemente anexados quedaron exentos de la ordenanza de *Villers-Cotterêts*. No obstante, en territorios más apartados como Nueva Francia (Canadá), Luisiana, Inglaterra, entre otros, había un mayor número de hablantes que en la misma Francia. Por otro lado, el francés se convirtió en la lengua de la diplomacia entre los países europeos, siendo una vez más la lengua franca de la época.

### ***El francés moderno (siglo XVII a la actualidad)***

#### ***La Ilustración***

Los historiadores franceses han ubicado la muerte del rey Luis XIV en el año de 1715, y con esta el inicio de la era de la Ilustración, periodo en el cual se desarrolló un gran movimiento intelectual en Europa y otras partes del mundo, caracterizado por el replanteamiento de todo tipo de creencias, desde la discusión por las ciencias o pseudociencias hasta el cuestionamiento de los derechos del pueblo y la monarquía. Los pensadores de la época estaban en contra de las supersticiones, la irracionalidad y la tiranía. Después de la muerte del rey Luis XIV sube al trono el rey Luis XV, inicialmente llamado el Bien amado, aunque en sus últimos años de reinado fue odiado por el pueblo debido a sus numerosas amantes y conflictos con la reina, y sobre todo por haber dejado a Francia en una grave crisis económica tras sus fracasos militares. Luis XV subiría al trono a la edad de cinco años con lo que se le dejó el gobierno al cardenal *André Fleury* hasta que Luis XV alcanzó la mayoría de edad. Al llegar a esta, el rey decidió casarse con *María Leszczyńska* hija del destronado rey de Polonia a la que el rey quiso mucho en un principio, hasta que perdió poco

---

<sup>28</sup> El término *patois* surgió precisamente en el siglo XVII, proveniente de la palabra *patte* o pata, porque según la gente culta los campesinos y la gente inculta hablaba con las patas, es decir tenían que hacer uso de gestos para poderse comunicar.

a poco su atractivo físico. El carácter del rey era débil y no le importaba mucho la política del reino, pues a pesar de haber llegado a la mayoría de edad le delegaba sus obligaciones al cardenal *Fleury*. A este último no le agradaba la presencia de la reina y en todo momento trataba de desviar la atención del rey, a veces presentándole mujeres que no lo influenciaran mucho. Cuando *Fleury* murió, el rey por primera vez tenía la libertad de gobernar, sin embargo, tomó decisiones equivocadas por su inexperiencia, sobre todo en el ámbito militar con las derrotas en las guerras contra Austria, Prusia e Inglaterra, lo que llevaría al reino en una grave crisis económica. En cuanto a las lenguas habladas en esa época, en algunas provincias al norte de Francia se hablaba el francés, aunque la mayoría de la gente lo hablaba usando el argot y provincialismos. Debido a la construcción de numerosos caminos a lo largo de esa región, el francés cada vez se expandía más, y en consecuencia un mayor número de personas lo empezaron a hablar. Al sur de Francia, en las regiones más recientemente anexadas, no se hablaba francés ni siquiera por la aristocracia, solo se hablaban las lenguas regionales (esta lengua sólo se hablaba durante los actos religiosos, por lo cual era llamada “la lengua del domingo”). En el resto de Francia, la gente era bilingüe pues hablaba francés y alguna lengua local. Aunque el francés seguía extendiéndose por el reino, los religiosos no estaban de acuerdo a que la gente lo aprendiera, debido a que los libros religiosos estaban escritos en latín o alguna lengua regional y el estado argumentaba que la enseñanza de este idioma a los campesinos era innecesaria, pues estos empezarían a leer libros distintos de los religiosos con lo que pretenderían buscar carreras profesionales abandonando la agricultura<sup>29</sup>.

Pese a que en el siglo anterior a la Ilustración el francés alcanzó su mayor pureza según los gramáticos, en la Ilustración se volvieron a introducir nuevas palabras. Entre estas se pueden identificar préstamos lingüísticos de lenguas de los reinos vecinos (alemán, castellano, italiano y sobre todo del inglés debido a sus relaciones constantes con Inglaterra). Además, se incluyeron términos técnicos tomados del latín y del griego. Si bien los censores de la lengua permitieron la adición de estas palabras extranjeras, no permitieron las palabras de uso común en Francia en lenguas *patois*, porque querían mantener la “pureza” del francés.

Por otro lado, Francia perdió el poder que tenía en Europa cuando el Reino Unido conquistó Canadá en 1760, con lo que el Reino Unido se convertiría en la nación más

---

<sup>29</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

influyente de Europa y el mundo. Desde entonces y hasta la fecha<sup>30</sup>, el inglés es la lengua franca a nivel mundial.

### ***El francés de la Revolución (1789-1870)***

El periodo de la Revolución está caracterizado por movimientos sociales originados en la Ilustración, cuya finalidad era conseguir cambios en la estructura social y política en Francia, partiendo de la abolición de la monarquía.

La monarquía empezó a tener cada vez mayor poder sobre su pueblo a partir del reinado de Enrique IV, el Buen Rey; durante el reinado de Luis XIII el cardenal *Richelieu* creó el Estado Centralizado y el rey Luis XIV, quien gobernó por 72 años, creó un régimen absolutista. Sin embargo, la popularidad de la monarquía fue en declive con Luis XV, que como se mencionó anteriormente, llevó a Francia a una grave crisis financiera, y con el reinado de Luis XVI la situación de la monarquía no mejoraría. Francia era una de las naciones más ricas de Europa, debido a su sistema de impuestos y cuotas aplicados a los campesinos, quienes representaban aproximadamente al 80% de la población. Sin embargo, llevaban una calidad de vida paupérrima, mientras que la nobleza vivía en la opulencia y la iglesia poseía el 10% de las tierras más productivas del territorio. A pesar de que la clase media poseía la mayor parte de la riqueza de la nación, no podían ejercer en el ámbito de la política. Debido a la crisis por la que pasaba Francia, tras la victoria contra el Reino Unido, este paradójicamente le arrebató Canadá con lo que el Reino Unido se convertiría en la nueva potencia mundial. Una de las consecuencias de esto fue el cambio del francés al inglés como lengua franca, la cual mantiene su estatus hasta la actualidad<sup>31</sup>. Por estas razones, la mayor parte del pueblo estaba harta de la monarquía y en consecuencia se fue gestando la Revolución Francesa (1789 a 1799). Al abolirse la monarquía, la gente de la clase media empezaría a tomar las decisiones de la nación, lo que incluye la imposición de su forma de hablar el francés, con lo que este dejaría ser de los monarcas para convertirse en la lengua de Francia. Así el francés se convertiría en un instrumento para crear una nación unida e

---

<sup>30</sup> Después de la Segunda Guerra Mundial, Estados Unidos de América es la nación más influyente del mundo con lo cual el inglés se sigue utilizando como lengua franca (lengua universal).

<sup>31</sup> El inglés sigue siendo la lengua franca a nivel mundial, después de que Estados Unidos demostró ser una potencia bélica con los ataques contra Japón, con lo que se terminaría la Segunda Guerra Mundial.

indivisible, bajo el lema “Libertad, Igualdad, Fraternidad”<sup>32</sup>, por lo que las lenguas regionales supondrían un obstáculo para propagar las ideas de los revolucionarios. Poco después, el 14 de junio de 1794, *Henri-Baptiste Grégoire* publicó su informe llamado *Rapport sur la nécessité et les moyens d'anéantir les patois et d'universaliser l'usage de la langue française* en donde da a conocer la situación lingüística de Francia de aquella época: se hablaban 33 lenguas *patois* o locales ininteligibles entre sí habladas por la mayoría de la población, mientras que el francés era hablado solo en 15 de 83 departamentos franceses, con alrededor de 3 millones de habitantes de los 25 millones de franceses, los cuales tenían diferentes sociolectos. Además, clasificó a las lenguas locales de forma errada y prejuiciosa. Al hablar del corso y el alsaciano comentó que eran formas muy degeneradas del italiano y del alemán respectivamente<sup>33</sup>. Finalmente, propuso imponer al francés parisino<sup>34</sup> como lengua de la nación. En la actualidad se trata de una lengua muy homogénea casi carente de dialectos y sociolectos. Más tarde, el 20 de julio de 1794, se llevó a cabo el “régimen del terror” (lingüístico) mediante un decreto, que condensaba la ideología de los revolucionarios, con las siguientes imposiciones:

- Sección 1.

“A partir de la fecha de la publicación de esta ley, ningún documento público podrá ser escrito en otra lengua distinta del francés en cualquier lugar de la República”.

- Sección 2.

“El mes siguiente a la fecha de la publicación de esta ley, ningún documento podrá ser registrado si no está escrito en francés, incluso los privados”.

- Sección 3.

---

<sup>32</sup> Este lema nació durante el movimiento de la Revolución y en la Constitución de 1848 aparece como un principio de la República.

<sup>33</sup> Abalain, H. (2007) *Le français et les langues historiques de la France*

<sup>34</sup> El francés parisino se caracteriza por la pronunciación de erres guturales y vocales nasales, por lo que al ser impuesto como estándar, esta pronunciación es muy característica del francés actual.

“Todo servidor público quien, a partir de la publicación de esta ley, use en cualquier documento lenguas distintas del francés será llevado ante los tribunales, condenado a seis meses de prisión y destituido”

- Sección 4.

“Esta misma sentencia será impuesta a cualquier registrador, el cual el mes siguiente de la fecha de publicación de esta ley, registre documentos en lenguas distintas del francés, incluso los privados”<sup>35</sup>

Este decreto no solo se limitaba a Francia, sino que también a sus colonias se les obligó a acatar estas reglas.

Sin embargo, había muchos motivos por los cuales el régimen del terror no tuvo el éxito que se esperó en contra de la “Torre de Babel” dialéctica. Los principales opositores fueron todas las personas involucradas con la iglesia, debido a que los libros religiosos estaban escritos en su mayoría en latín o en alguna lengua regional, con lo que se dificultaría el adoctrinaje. Otro factor del fracaso momentáneo del régimen se debió a la falta de fondos monetarios para poder pagar a los suficientes profesores que pudieran dar clases en cada villa en el territorio francés por pequeña que fuera (incluso en París no se contó con los recursos suficientes).

### ***El régimen dictatorial de Napoleón (1799-1815)***

Napoleón Bonaparte nació el 15 de agosto de 1769 en Ajaccio, la capital de la isla de Córcega, donde se habla el corso -una lengua emparentada con el italiano-, y naturalmente esta fue la lengua materna de Napoleón. Poco después del fin de la Revolución francesa, el 11 de noviembre de 1799, Napoleón se convirtió en el primer cónsul de la República y el 18 de mayo de 1804, Emperador de los franceses, hasta su derrota en manos de los belgas en la batalla de *Waterloo* el 18 de junio de 1815. Su régimen fue muy estricto pues su país estaba en crisis económica y bajo la anarquía. Para instaurar el orden y mejorar la economía de la República, creó una dictadura tomando el control absoluto, con lo cual estimuló la economía y mejoró las comunicaciones, a costo de mantener a la nación en un estado de guerra

---

<sup>35</sup> Traducción propia de texto en [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

permanente. Al ser un hablante de corso no estaba muy de acuerdo con las políticas en contra de los dialectos y con el pretexto de mejorar la economía de la nación encomendó la educación a la iglesia para ahorrarse el salario de los profesores que daban clases en francés, con lo cual la iglesia volvía a dar clases en las distintas lenguas regionales o en latín.

Mientras, se volvió a reestablecer la *Académie française* después de que se suspendiera por la Revolución, durante este periodo el francés no sufrió grandes cambios, excepto en el vocabulario, debido a que la nación estaba en guerra constante contra otras naciones para anexarlas al imperio, se tomaron nuevamente prestamos del inglés. En 1823 la *Académie française* publicó la Gramática francesa de *François Noël* y *Jean-Pierre Chapsal* cuyo título original fue: *Nouvelle Grammaire française sur un plan très méthodique, avec de nombreux exercices d'orthographe, de syntaxe et de ponctuation, tirés de nos meilleurs auteurs, et distribués dans l'ordre des règles*. Esta gramática tuvo ocho ediciones en Francia hasta 1889 y otras más se hicieron en Montreal, Canadá y una traducción en Filadelfia en 1878. Este libro era esencial para la educación de los niños de Francia, y de algunas colonias, como Canadá, y se caracterizó por ser una gramática inflexible que no admitió los cambios naturales que cualquier lengua experimenta, de esta manera la gente con un mayor poder adquisitivo se apegaba a la manera de escribir y hablar que dictaba la gramática, mientras que la gente con menor poder adquisitivo no se esforzaba en seguir las reglas de esta, por lo que el uso de una buena ortografía representó un símbolo de distinción social. Por otro lado, gracias a la mejora de las comunicaciones y transportes se pudo extender de manera más fácil el uso del francés parisino, el cual ya se había establecido al término de la Revolución como el estándar en toda la República.

### ***El Segundo Imperio (1852-1870)***

El 2 de diciembre de 1852, Napoleón III proclama el nuevo Imperio bajo su mando, ya que había sido el único presidente de la Segunda República desde 1848. Este régimen dictatorial se caracterizó por la opresión a la libertad de expresión, controlando a la prensa y a sus detractores; además tomó parte de una política exterior agresiva. A pesar de lo anterior, no tenía muchos opositores gracias a la buena situación económica nacional y a su carácter de gobernante conservador. Su régimen finalizó tras su captura en Sedán, Francia mientras combatía contra el ejército prusiano el 1 de septiembre de 1870 y sólo dos días después se

proclamaría la Tercera República vigente hasta la actualidad, donde se conseguiría estabilizar Francia. Durante este tiempo, el francés se vio nuevamente enriquecido con nuevo vocabulario, donde no solo aparecen palabras creadas por los anarquistas y opositores liberales, sino que también se agregaron palabras de las ciencias, en especial de la física, la química, y la medicina; además de las nuevas ciencias como la arqueología, la paleontología, la zoología, la lingüística, entre otras. También se agregaron los nombres de nuevos inventos y medios de transporte, por ejemplo, *locomotiv* (locomotora), *automobile* (automóvil), *omnibus*, (ómnibus), *tram* (tranvía), *autobus* (autobús), etc. Por último, a través de los medios de comunicación impresos se propagaron diversos neologismos creados por los recientes anuncios publicitarios. Todas estas palabras se imprimieron en un diccionario creado por Maximilien-Paul-émile-Littré y Pierre Larousse. El francés era hablado en todo el territorio (incluidas sus colonias) y las lenguas regionales se empezaban a influenciar con la inclusión de palabras del francés.

### ***Francés contemporáneo (1870- )***

Con la caída del segundo imperio surgió la tercera república. Durante esta transición se desarrollaron cambios en cuestiones políticas, sociales, tecnológicas, científicas, económicas, e incluso de los deportes. Se introdujeron de nuevo distintas palabras a causa de estos cambios y continuó la adopción de préstamos del inglés gracias a la cercanía con Inglaterra y sus relaciones con esta nación. A pesar de que el francés ya se hablaba en todo el territorio y que a partir de la caída del Segundo Imperio se empezó a enseñar en toda la República, la unidad lingüística tardaría décadas en poderse consolidar en Francia (caso contrario de Canadá donde se había unificado en sus antiguas colonias desde hace tiempo). Entre los años de 1881 y 1882, Jules Ferry redactó una serie de leyes en las que se establecía que la educación primaria debía ser obligatoria, laica y gratuita para los niños y las niñas, con lo que se excluía a la iglesia de impartirla, de esta manera el francés volvería a ser la lengua para la impartición de la educación. Durante el siglo XX hasta la época de 1960 se habían escrito más de 40 leyes sobre educación, prensa y administración que promovían el uso del francés. Poco después se intensificarían los sentimientos anti-dialecticos, por ejemplo, en 1972 el entonces presidente Georges Pompidou declaró: “No hay lugar para lenguas y culturas regionales en una Francia que debe dejar su sello en Europa”. Décadas más tarde, el

13 de mayo de 1992, el diputado Robert Pandrau, durante los debates del Tratado de la Unión Europea celebrado en *Maastricht*, Países Bajos, declaró:

“Pago tributo a las escuelas seculares de la República, las cuales a menudo han impuesto el francés con autoridad considerable, algo que fue necesario, en cara de todas las fuerzas sociales e incluso del obscurantismo religioso que existió en la época. Estoy igualmente encantado que la televisión ha sido una fuerza unificadora para la lengua. Es tiempo para convertir el francés nuestra lengua. Si necesitamos enseñar a nuestros niños otra lengua, no permitamos desperdiciar su tiempo con dialectos que solo hablaran en sus villas. Permitamos enseñarles una lengua internacional tan pronto como sea posible”<sup>36</sup>.

El 4 de agosto de 1994 se decretó la ley Toubon, creada originalmente para frenar el uso excesivo de anglicismos en Francia. Esta ley hace obligatorio el francés en documentos, publicidad, comunicaciones, lugares de trabajo, software, manuales de usuario, etc. Esta ley además restringió el uso de otras lenguas como medios de enseñanza en las escuelas, negándoles el financiamiento en caso de hacerlo. Sin embargo, hubo leyes y personas que estaban a favor de las lenguas y dialectos regionales. La primera ley escrita a favor de algunas de estas lenguas fue la ley *Deixonne*, propuesta por Maurice Deixonne, promulgada el 13 de enero de 1951. Esta ley autoriza la enseñanza de ciertas lenguas regionales que son el vasco, bretón, catalán y occitano, anexándose el corso<sup>37</sup> a esta ley hasta el 16 de enero de 1974. En una revisión a la Constitución, el 23 de julio de 2008, se modificaron varios artículos y otros se crearon, entre los cuales está el artículo 75-1 que declara: “Las lenguas regionales pertenecen al patrimonio de Francia”<sup>38</sup>.

---

<sup>36</sup> Traducción propia de texto en [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

<sup>37</sup> Estas lenguas, junto con el provenzal, fueron utilizadas en el modelo de CGS propuesto. Es probable que gracias a la ley *Deixonne*, estas lenguas todavía no han desaparecido y se pueden encontrar textos en estas lenguas en internet.

<sup>38</sup> <https://www.legifrance.gouv.fr/affichTexteArticle.do?idArticle=LEGIARTI000019241104&cidTexte=LEGI TEXT000006071194&dateTexte=20090810> Abecassis M., Ledegen G., Zouaoui K. (2011) La francophonie ou l'éloge de la diversité.

### 4.1.2 *Lenguas francesas*

#### *Occitano*

El aquitano es una lengua hablada por unas 218,310 personas en el “Mediodía francés” (territorio al sur del río Loira, Francia). Como se mencionó anteriormente, el nombre de esta lengua deriva de la palabra *oc* que en occitano significa “sí”, así mismo esta palabra proviene del latín *hoc*. El poeta italiano Dante Alighieri hizo una clasificación de lenguas en su libro *De vulgari eloquentia* donde divide a las lenguas romances en tres grupos: lenguas de *oïl*, lenguas de *oc* y lenguas de *sì*. Las lenguas de *sì* eran habladas en Iberia e Italia, las lenguas de *oïl* al norte del río Loire en Francia y las lenguas de *oc* al sur del río Loire.<sup>39</sup>

El occitano surgió en Occitania, la zona que comprende el Mediodía francés, valle de Arán en Cataluña y los valles occitanos que pertenecen a la región italiana del Piamonte. Durante la Edad Media, Occitania se convirtió en uno de los lugares más importantes de la cultura romana, a partir del siglo IX, razón por la cual el occitano fue una de las primeras lenguas que sustituyeron al latín como lengua con la que se escribían los textos científicos, documentos oficiales y la literatura.

La edad de oro de la lengua occitana terminó con la muerte del rey Pedro II de Aragón y la derrota de las tropas occitano-aragonesas durante la Cruzada albigense, una guerra en contra de los cátaros, cuya religión tenía dos dioses uno benévolo y otro maléfico y que se situaron principalmente en *Languedoc*. Esta guerra fue comenzada por iniciativa del papa Inocencio III con la ayuda de la dinastía de los Capeto que eran los reyes de Francia. Con este conflicto se frenó el catarismo y la expansión del reino aragonés con lo que Francia se apoderó de Occitania.

Según la mayoría de los lingüistas y escritores se considera al occitano como una lengua de la cual se desprenden distintos dialectos, por lo que el occitano también se le considera una familia de lenguas, conformada por el limosín, auvernense, gascón, *languedoc*, provenzal y provenzal alpino. El catalán alguna vez fue un integrante más de la familia occitana considerándolo como provenzal o lemosín.

---

<sup>39</sup> <https://www.ethnologue.com/language/oci>

## **Provenzal**

El provenzal (*provençau* en provenzal) es considerado un dialecto del idioma occitano<sup>40</sup>, hablado en la antigua provincia de Provenza (situada al sureste de Francia y que se extiende desde la orilla oriental del río Rodano inferior, al oeste hasta la frontera con Italia y al este limita al sur con el mar Mediterráneo).

Entre los siglos XII y XIII se escribieron muchas obras en lenguas de *oc*, entre estas el provenzal. En un principio se escribieron textos poéticos y hacia los siglos XIV y XV se escribieron textos en prosa, así como textos legales, científicos y de temas lingüistas.

El reino francés se fue desarrollando con el paso de varios siglos bajo dinastías situadas en París, al norte de Francia, donde se hablaban lenguas de *Oïl* que eran utilizadas en todos los ámbitos, desde la literatura y los textos científicos hasta los textos legales. A finales del siglo XV, Luis XI amplió el territorio de Francia al anexar territorios del sur, quedando Provenza bajo su dominio en 1482 y en el año 1539 el rey Francisco I (bisnieto de Luis XI) promulga el decreto de *Villers-Cotterêts*, por lo cual todos los actos oficiales deben escribirse en francés, la lengua de la dinastía de los Capeto. A pesar de lo anterior, al este del río Rodano, se sigue utilizando el provenzal por la mayoría de la gente, mientras que el francés es poco utilizado, a excepción de los burgueses de Marsella.<sup>41</sup>

## **Catalán**

El catalán es una lengua romance occidental, hablada en Cataluña, Comunidad Valenciana, Islas Baleares, franja oriental de Aragón, la sierra El Carche (las regiones mencionadas anteriormente pertenecientes a España), Rosellón (Francia), Ciudad de Alguer (Italia) y como único idioma oficial en Andorra.

El catalán es hablado por 4,353,000 personas como lengua materna. Con un total de 9,229,420 hablantes, de los cuales 5,150,000 se sitúan en España, 30,000 en Andorra, 100,000 en Francia y 20,000 en Italia.<sup>42</sup>

El catalán es considerado como un idioma con características de transición, ya que

---

<sup>40</sup> <https://www.ethnologue.com/language/oci>

<sup>41</sup> <http://faculty.education.illinois.edu/westbury/paradigm/documents/Graves.pdf>

<sup>42</sup> <https://www.ethnologue.com/language/cat>

cuenta con elementos de la subfamilia iberorromance y de la familia galorromance, además es bastante próximo al occitano, por tal motivo los lingüistas consideran a este dentro de la subfamilia occitano-romance y como ocurre con la mayoría de las lenguas ha sido influenciada por otras habladas en España, Francia e Italia.

El catalán es usado de forma habitual por el 47% de los habitantes de Cataluña siendo el segundo idioma en uso en esta región seguida del castellano.

La Generalidad de Cataluña promueve el uso del catalán en ámbitos educativos y administrativos, así como en los medios de comunicación, no solo en Cataluña, sino que en todos los territorios de habla catalana aun fuera de España.

El catalán surgió a partir del latín vulgar, evolucionando independientemente de otras lenguas romances. En los siglos VII y VIII aparecen textos en los que se observan elementos propios del catalán. Posteriormente, en los siglos X y XI, se empezaron a escribir los primeros textos en catalán. En el año de 1028 se escribe el juramento feudal, en 1090 *Greuges de Caboet* -que son los primeros textos escritos totalmente en catalán- y en 1474 se imprime el primer libro en catalán y el primero en España con el título: *Obres e trabes en lahors de la Verge Maria*.

Algunos especialistas dicen que en realidad el catalán y el occitano son dos variedades de un idioma en común y forman el grupo occitano-romance<sup>43</sup>, e incluso al catalán se le considero como un dialecto del provenzal<sup>44</sup> que a su vez se le considera un dialecto del occitano.

Existen dos estándares de catalán: uno que está regulado por el *Institut d' Estudis Catalans*, que es el estándar más aceptado y que tiene como base la ortografía dada por el lingüista Pompeu Fabra y que el el más cercano al catalán hablado en Barcelona. El otro estándar es el regulado por la *Acadèmia Valenciana de la Llengua* que es utilizado por la Comunidad Valenciana y el Carche, este es más apegado al valenciano que es una de las variantes del catalán.

Como otras lenguas, el catalán posee algunos dialectos que son muy parecidos entre sí,

---

<sup>43</sup> [http://www.univ-montp3.fr/uoh/occitan/une\\_langue/co/module\\_L\\_occitan\\_une%20langue\\_7.html](http://www.univ-montp3.fr/uoh/occitan/une_langue/co/module_L_occitan_une%20langue_7.html)  
[http://cle.ens-lyon.fr/servlet/com.univ.collaboratif.util.LectureFichiergw?ID\\_FICHIER=1332154733834](http://cle.ens-lyon.fr/servlet/com.univ.collaboratif.util.LectureFichiergw?ID_FICHIER=1332154733834)

<sup>44</sup> Diez F. (1874) "Grammaire des Langues Romanes", Tome 1.

estos se dividen en dos grupos: el bloque dialectal oriental (donde están incluidos los dialectos insulares, central y de Francia) y el bloque dialectal occidental (al cual pertenecen el valenciano y el noroccidental).

### **Vasco**

El idioma vasco o euskera (*Euskal* en vasco) es una de las pocas lenguas aisladas en Europa y la única de este tipo de Europa Occidental. Se le llama aislada porque no desciende del protoindoeuropeo, a pesar de esto influyó al castellano en cuanto al sistema fonético pentavocálico.

Desde el periodo de la Baja Edad Media hasta el siglo XIX cada vez menos personas lo hablan, debido a que al ser una lengua aislada la gente prefería cada vez hablar alguna lengua romance para facilitar el comercio con otros pueblos. Pero a finales de 1950 se comenzó el trabajo de estandarización para evitar su extinción, esto con el fin de no tener muchos dialectos para mejorar la comunicación entre individuos hablantes de euskera con el fin de incentivar su uso y conservar su identidad como pueblo.

A partir del fin del régimen dictatorial, con la muerte de Francisco Franco, se creó la constitución de 1978, que entre otras cosas les concedía a las comunidades autónomas el derecho de declarar como oficiales sus propias lenguas<sup>45</sup>, lo que se plasmó para el País Vasco, en el *Estatuto de Guernica* al siguiente año. Sin embargo, en el País Vasco francés, el vasco (como otras lenguas), no es reconocido por el gobierno francés y simplemente es llamado otra lengua *Patois* (una lengua no oficial en Francia, término usado despectivamente). El antecesor del vasco fue el euskara arcaico o aquitano, este nombre se debe a que se hablaba en la región de Aquitania en Francia.

### **Bretón**

El bretón (o *Brezhoneg* en bretón) es una lengua celta y es una de las lenguas británicas (junto al galés) que se siguen hablando actualmente. Las lenguas británicas pertenecen a las lenguas insulares junto con las lenguas goidélicas (irlandés, gaélico escocés y manes). Tanto las lenguas británicas como goidélicas pertenecen al protocelta, que a su vez desciende del

---

<sup>45</sup> [http://administracion.gob.es/pag\\_Home/espanaAdmon/comoSeOrganizaEstado/ComunidadesAutonomas.html](http://administracion.gob.es/pag_Home/espanaAdmon/comoSeOrganizaEstado/ComunidadesAutonomas.html)

protoindoeuropeo. La palabra britónica proviene de la palabra inglesa *Brythonic*, que a su vez se deriva de la palabra galesa *Brithon*, que significa “indígena celta”. Estas lenguas surgieron en Gran Bretaña desde la Edad de Hierro (aproximadamente en el siglo XII a.C.), a pesar de este hecho el bretón actualmente es hablado en Bretaña, Francia (que es una península al noroeste de Francia entre el Canal de la Mancha y el Golfo de Vizcaya) debido a las migraciones de los britanos antes del inicio de la Edad Media, quienes provenían de Gales y Cornualles (sur de Inglaterra), hacia Armónica (actualmente Bretaña) a causa de las Invasiones anglosajonas.

El desarrollo del bretón se divide en 4 etapas:

- Bretón primitivo (500-600).
- Breton antiguo (600-1000).
- Bretón medio (1000-1600, etapa en la cual adopta una gran cantidad de préstamos lingüísticos del francés).
- Bretón moderno (1600-actualidad).

Las leyes francesas en los siglos XIX y XX hacían obligatoria la educación exclusivamente en francés prohibiendo el uso de las lenguas *Patois* para la enseñanza e incluso en los medios de comunicación, esto con el fin de que cada vez más franceses dejaran de hablar lenguas de este tipo y empezaran a hablar francés, con lo cual se pretendía un mayor nacionalismo.

Durante la década de 1970 se reabrieron escuelas de bretón y las autoridades de Bretaña promovieron su uso debido a que solo la gente mayor hablaba bretón mientras los niños solo aprendían a hablar francés, por ser esta la lengua que se enseñaba en las escuelas, y por la humillación a la que se exponían si hablaban en bretón. Se trata, por tanto, de una lengua en peligro de extinción, hablada aproximadamente por 206,000 personas en Bretaña en la región denominada como Baja Bretaña, aunque también se habla en la Alta Bretaña junto con el francés y galo (ambas lenguas de *Oïl*). Este idioma llegó al Continente Americano, específicamente a Argentina, Canadá y Estados Unidos, debido a las migraciones entre las décadas de 1920 a 1950.

Actualmente existen cuatro dialectos del bretón:

- El veinnetés (*Gwendeg*).
- El cornuallés (*Kerneveg*).
- El leonés.
- El tregorés.

Además, el bretón cuenta con dos sistemas de ortografía, a pesar de los esfuerzos de unificación. Uno de ellos es el *Zedacheg* (ZH) y el otro es el *Orthographe Universitaire* (OU). El *Zedacheg* se estableció en 1941 para alojar a los cuatro dialectos mencionados. La “Ortografía Universitaria” fue propuesta en 1955 por *Falc’hun*. Esta ortografía fue reconocida por el gobierno francés ya que promovía una pronunciación y conjunto de símbolos más cercanos al francés. Por estos motivos la mayoría de la gente se rehusó a adoptarla, prefiriendo al *Zedacheg*.

### **Corso**

El idioma corso es una lengua romance perteneciente a la subfamilia italo-dálmata y al grupo toscano (es un idioma muy parecido al italiano), hablado por más de 100,000 personas, principalmente en la isla de Córcega, Francia, donde es cooficial con el francés. Esta lengua también es muy utilizada en la isla vecina de Cerdeña, Italia.

Córcega fue propiedad de varios reinos y repúblicas; desde los pueblos vándalos, ostrogodos, bizantinos, lombardos. Luego pasó a ser concedido al Papa Esteban II por parte de Pipino el Breve (rey de los francos y padre de Carlomagno) tras haberle pedido ayuda para sacar a los lombardos de Córcega, pues eran considerados una amenaza para Roma. En el siglo XI Córcega fue liberada de los árabes por los ejércitos de las Repúblicas de Pisa y Génova. A partir de este hecho, Córcega sería anexada a la República de Pisa, entonces empezó una gran inmigración proveniente de La Toscana, por lo que dos tercios de la isla de Córcega quedarían influenciados con su cultura, su lengua y su relación con otros pueblos de la Italia actual. En 1284, la República de Génova gana la Batalla de Meloria contra la República de Pisa, con lo cual (entre otras cosas) la República de Génova<sup>46</sup> toma el control

---

<sup>46</sup> La República de Génova se extendía por toda la región de la actual Liguria, Italia que colinda con las zonas

de la isla de Córcega.

### **Francés**

- El francés es una lengua practicada por 229,217,640 personas en todo el mundo, de estas, 76 millones lo aprendieron como lengua materna y 153 millones como segunda lengua.
- En Francia es la lengua materna de 60 millones de personas y la segunda lengua de 2,900,000.
- En Bélgica es uno de los idiomas oficiales junto con el neerlandés y el alemán, es el segundo idioma más importante (seguido del neerlandés) es empleada como lengua materna por aproximadamente 4 millones de personas y por 5 millones de personas como segunda lengua.
- En Suiza es oficial junto con el alemán, italiano y romanche; es la segunda lengua más usada (seguida del alemán) es la lengua materna de casi 2 millones de personas y la segunda lengua de 3.5 millones.
- En el continente americano, el mayor país francófono es Canadá, donde es cooficial junto con el inglés, hablado por 7.3 millones de personas como lengua materna y 2.7 millones como segunda lengua.
- El francés es una lengua utilizada en los 5 continentes, ya que antes de la segunda guerra mundial, era una lengua franca debido a distintas invasiones y conquistas en el mundo.
- Existen distintas variedades o dialectos del francés De estas, las que son habladas en el continente americano son llamadas criollo francés y son utilizadas en las Antillas, Guyana Francesa, Florida (EEUU), *Louisiana* (EEUU) y Haití.<sup>47</sup>

---

de Francia donde se habla el provenzal. La República de Génova se anexó al Reino de Cerdeña en 1815, a su vez, el Reino de Cerdeña fue el reino principal que fundaría al Reino de Italia conquistando y anexando nuevos territorios durante el siglo XIX hasta formar la Italia actual.

<sup>47</sup> <https://www.ethnologue.com/language/fra>

## 4.2 Familia Germánica.

La familia lingüística germánica cuenta con solo doce lenguas modernas, de las más de 5,000 lenguas practicadas en el mundo en la actualidad, y entre cuarenta y cincuenta lenguas mixtas (criollas). Pese a esto, el número de hablantes de lenguas germánicas más utilizadas es de aproximadamente 1,117 millones (aproximadamente el 15% de la población mundial), de los cuales 451,143,150 son hablantes nativos<sup>48</sup>. Además, algunas de las lenguas de esta familia se han disperso por el mundo y son muy influyentes, siendo uno de sus más notables representantes el inglés. En Internet, en cantidad de páginas disponibles, se encuentran en primer lugar el inglés, y después el alemán.<sup>49</sup>

¿Cómo es posible que diversas lenguas se inflencien entre sí?

Cada persona tiene una lengua madre, con la cual se puede comunicar con los miembros de su misma comunidad, además posee su propio sociolecto que lo puede identificar en una zona o nivel sociocultural más específico. Sin embargo, los individuos se desplazan más allá de sus propias regiones e incluso países debido a distintas razones, por ejemplo, migraciones, invasiones, etc. Las migraciones también tienen un papel importante en las variaciones de una lengua, debido al intercambio cultural y lingüístico que se puede generar. Como en el caso del inglés de Estados Unidos que ha sido influenciado por los inmigrantes latinoamericanos, especialmente de México. Otro ejemplo es, la palabra de inglés antiguo *æppel* significaba “cualquier tipo de fruta”, pero al emigrar a otras regiones, se tomó la palabra latina *fructus* para denominar genéricamente a cualquier fruto y la palabra *æppel*, de la cual derivan *apple* en inglés y *Apfel* en alemán, tomó el significado de un fruto específico: la manzana. Para finalizar con estos ejemplos, en 1066 William el Conquistador, que hablaba normando (un dialecto del francés antiguo) tomó posesión de la isla en la que actualmente se encuentran Inglaterra y Escocia, fue una fuerte influencia para el inglés, debido a que la nobleza hablaba normando con lo que este se volvió la lengua oficial de las Islas Británicas hasta el acenso al trono de Henry IV quien hablaba inglés medio. De esta manera el 70% de las palabras en inglés moderno tienen su origen del normando.

---

<sup>48</sup> <https://www.ethnologue.com>

<sup>49</sup> <http://unesdoc.unesco.org/images/0018/001870/187016e.pdf>

Otro caso de influencia entre lenguas se da con las colonizaciones de ciertas regiones donde ya había habitantes con su propio idioma como el caso del actual territorio mexicano donde existían varias civilizaciones antes de la llegada de los españoles, entre estas destacan la azteca y la maya de donde se obtuvieron préstamos para el español y otras lenguas. El criollo haitiano se considera un dialecto del francés, que surgió por la influencia de lenguas africanas como el *wolof*, *fon* y *éwé*<sup>50</sup>. En la actualidad, gracias al fenómeno de la globalización, económica y de la información las lenguas más habladas en el mundo, sobre todo el inglés, están influenciando a otras en diversos países.

Hoy en día, el inglés es un idioma considerado [por distintas organizaciones a nivel internacional entre estas, la Organización de las Naciones Unidas (ONU)] *lingua franca*<sup>51</sup>, es decir, que facilita la comunicación entre los individuos de diversas naciones, incluso considerándolo una lengua universal. El inglés se utiliza en comercio, divulgación científica y técnica, educación, eventos y espectáculos internacionales, turismo, tratados internacionales e incluso como medio de socialización.

Otra lengua muy importante de esta familia es el alemán, única lengua oficial de Alemania y Austria. Además, es una de las lenguas oficiales en Suiza (aproximadamente el 64% de su población lo usa), Liechtenstein, Luxemburgo (junto con el francés y el luxemburgués, que es otra lengua germánica) y Bélgica (junto con el francés y el neerlandés). Esta lengua tiene numerosas variedades locales, que pese a llamarse en su conjunto alemán, resultan poco o nulamente inteligibles entre hablantes de distintas variantes.

El alemán es el idioma más practicado en la Unión Europea, siendo la lengua materna de aproximadamente 77 millones de personas y es la segunda lengua germánica más importante en el mundo con alrededor de 53 millones de usuarios.<sup>52</sup>

El tercer idioma germánico en cuanto a representación es el neerlandés (holandés<sup>53</sup>). Otras variantes o dialectos del neerlandés son el flamenco, el limburgués y el afrikáans. Este

---

<sup>50</sup> <https://www.ethnologue.com/language/hat>

<sup>51</sup> <http://www.un.org/en/events/englishlanguageday/>

<sup>52</sup> <https://www.ethnologue.com/language/deu>

<sup>53</sup> La Real Academia de la Lengua Española indica que pueden utilizarse la palabra holandés para referirse al neerlandés, aunque en ámbitos más formales se deben diferenciar.

último muchas veces se le ha considerado una lengua criolla distinta del neerlandés nacida en África.

La familia germánica surgió en una zona que comprende el norte de Alemania y parte de Polonia, en las desembocaduras de los ríos Elba, Óder y Vístula. Esta familia se desprendió del indoeuropeo, formando la lengua protogermánica. Los registros más antiguos del protogermánico fueron inscripciones rúnicas escandinavas de principios del siglo III.<sup>54</sup> La lengua germánica más antigua de la que se tenga registro es el gótico. Y el escrito más antiguo de este idioma es una traducción de la biblia, la cual utilizó grafías basadas en el alfabeto griego con influencias del latín.

La familia germánica se divide en tres grupos: germánico oriental, nórdico y occidental, en ocasiones también llamado germánico del sur.

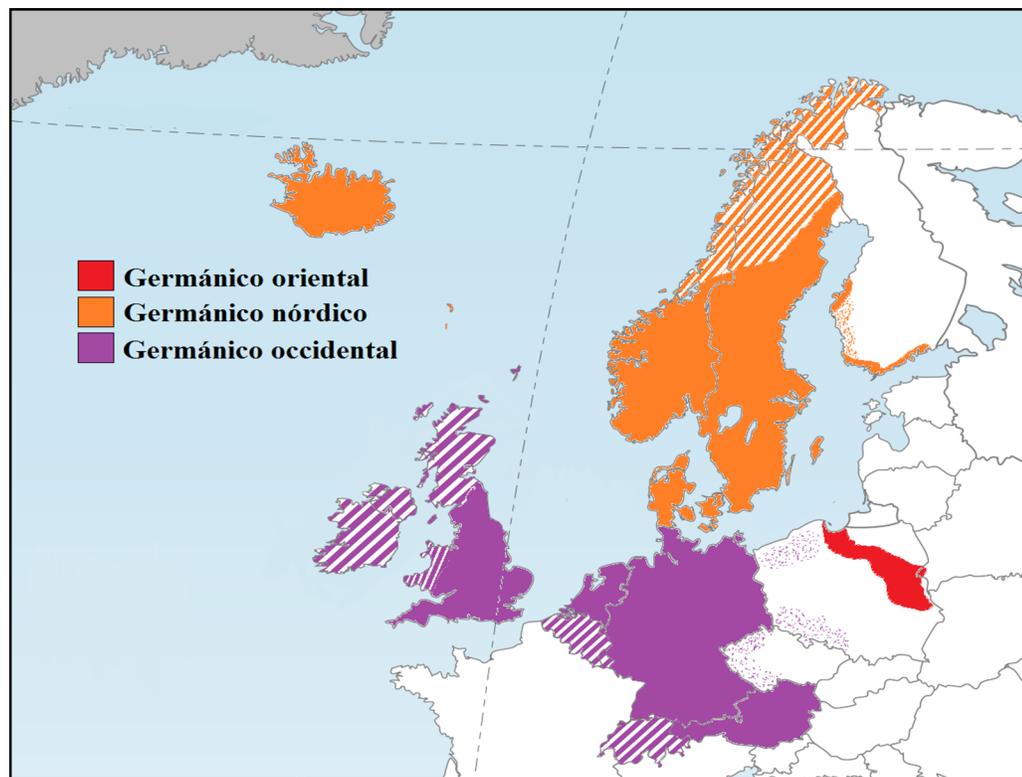


Figura 5. Distribución geográfica de las lenguas germánicas

<sup>54</sup> <https://global.britannica.com/topic/Germanic-languages>

El germánico oriental sólo cuenta con el extinto lenguaje gótico como representante. El germánico occidental tiene los siguientes: alemán, inglés, yidis, flamenco, neerlandés, afrikáans, *Plattdeutsch* (bajo alemán) y alemán de Pensilvania. Las lenguas que conforman al germánico nórdico son el noruego, islandés, danés, sueco y feroés.

El origen de las lenguas germánicas orientales (también llamadas del este) está relacionado con la migración de tribus bárbaras. Al final de la era precristiana existieron varias tribus germánicas como los vándalos, burgundios y godos, que dejaron su lugar de origen, sin embargo, fueron solo los godos quienes dejaron registros escritos. Algunos miembros de esta tribu migraron hacia el mar Báltico, específicamente hacia la costa este del río Óder. Algunos otros fueron hacia los Balcanes y de ahí volvieron a migrar hacia el oeste a tierras que actualmente se localizan en Italia, Francia y España. Debido a que las primeras migraciones de los góticos fueron hacia el este a su lengua se le califica, como ya se mencionó, germánico del este.<sup>55</sup>

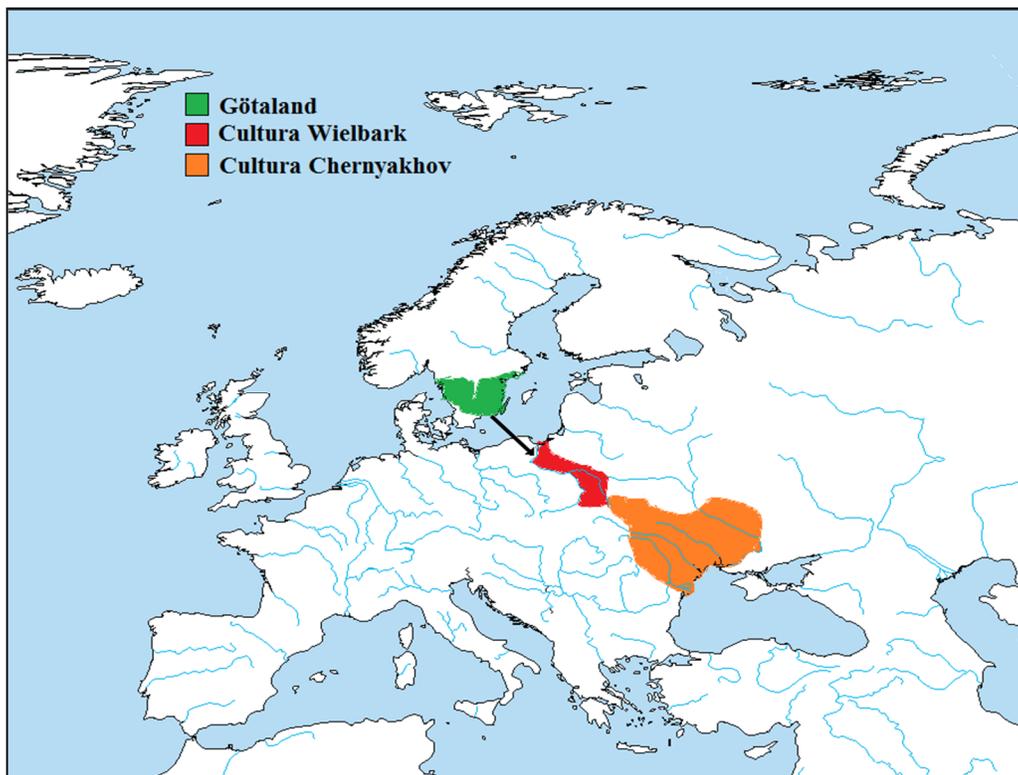


Figura 6. Distribución de los Godos

<sup>55</sup> <https://global.britannica.com/topic/Germanic-languages>

Fue aproximadamente entre los años 800 a 1050 d.C. que los hablantes del germánico nórdico se asentaron en lugares tan remotos como Islandia, Groenlandia, las islas Feroe y *Shetland* (isla de Escocia en el Mar del Norte). Esto está documentado en las sagas vikingas, en especial la *Saga de Erik el Rojo* (*Eiríks saga rauða*<sup>56</sup>).

En cuanto al germánico occidental, este quedó definido a partir de las invasiones de tribus anglos y sajonas a Inglaterra durante los siglos V y VI, creando después de su contacto con el latín y el normando, la lengua inglesa actual.

A partir de la llegada de Cristóbal Colón al continente americano en el siglo XV, los europeos empezaron a colonizar diversas zonas de este. Naciones como el Reino Unido y Países Bajos llegaron a Estados Unidos, Canadá, Guyana, Surinam, entre otros e impusieron sus idiomas. Más tarde, durante el siglo XIX, hubo un auge expansionista por parte de los europeos. El Reino Unido, Alemania, y Bélgica invadirían y conquistarían estados en África<sup>57</sup>, Asia y Oceanía.

El yidis apareció a partir de las migraciones de judíos (que hablaban alemán) hacia tierras eslavas entre los siglos XII y XVI, de ahí esta nueva lengua se difundió hacia Europa, Palestina-Israel, América, Sudáfrica y Australia durante los siglos XIX y XX.

Otro factor que influyó para que el protogermánico se dividiera en una docena de lenguas diferentes fue la estandarización. A pesar de que algunas de las lenguas germánicas se parecen y se entienden entre sí, por ejemplo el noruego (*bokmål*) y el danés se consideran lenguas distintas<sup>58</sup>, mientras que, el alto y bajo alemán se consideran como dialectos de la misma lengua. Esto se debe a la estandarización que el gobierno alemán promueve, la cual impone uniformidad en ésta lengua.

---

<sup>56</sup> [http://www.sagadb.org/eirik\\_saga\\_rauda.en](http://www.sagadb.org/eirik_saga_rauda.en)

<sup>57</sup> Entre los años 1884 y 1885, Alemania, Francia y Reino Unido convocaron a la conferencia de Berlín a varios países (que en su mayoría eran europeos) para poner fin a las disputas por los territorios del continente africano. Entre los países de habla germánica se encontraron a Alemania (que conquistó Camerún, Burundi, Namibia, etc.), Bélgica (República del Congo) y el Reino Unido (Egipto, Kenia, Nigeria, entre otros). [Moreno J, 1976]

<sup>58</sup> El noruego *Bokmål* en realidad es una lengua creada con base al danés con vocabulario y gramática similar.

La estandarización de las lenguas se realiza para unificar los idiomas o dialectos que se hablan en un territorio con una misma cultura y economía con el fin de que sus pobladores se puedan comunicar de manera más eficiente y así tener una identidad nacional más fuerte. Por las razones anteriores se crearon los estándares para el danés, sueco, alemán, neerlandés e inglés.

Durante el siglo XX Noruega e Islandia se independizaron (de Suecia en 1905 y de Dinamarca en 1944 respectivamente) por lo que se crearon los estándares del noruego e islandés (proveniente del nórdico antiguo) con la finalidad de crear un sentimiento nacionalista en sus pobladores.

Mientras que el idioma feroés, hablado en las Islas Feroe, es una lengua que evolucionó del nórdico antiguo. Aunque estas islas son un país autónomo del Reino de Dinamarca, y por consiguiente se habla danés, no han dejado de usar su lengua original. De esta manera, utilizan como lenguas el feroés y el danés.

El afrikáans es un dialecto del neerlandés que surgió en el extremo sur del continente africano, debido a la colonización de esa región por los neerlandeses (provenientes de la Unión de Provincias Unidas, actualmente Países Bajos) fundando la Colonia del Cabo cuya lengua oficial sería el neerlandés, el cual sería influenciado por las lenguas locales, el inglés y el portugués. Actualmente el afrikáans es hablado en Sudáfrica, Namibia y Botsuana.

Noruega cuenta con dos estándares de escritura: el *Bokmål*, o lengua del libro, es el más hablado por sus habitantes (85 a 90% de la población) y está basado en el danés. El segundo estándar es el *Nynorsk*, noruego nuevo, que es usado como lengua escrita por el resto de la población; este estándar es una mezcla de varios dialectos noruegos.

#### **4.2.1. Germánico Nórdico**

Dinamarca, Suecia, Noruega e Islandia son naciones vecinas, con un pasado en común. Estos territorios formaron parte de lo que se llamó Escandinavia, en el periodo entre los siglos X y XIX. Existieron pueblos de navegantes y guerreros, los vikingos, que además de viajar hacia las naciones del Oeste como Inglaterra, Escocia y Normandía también alcanzaron

lugares como Islandia, Groenlandia y las islas Feroe, esparciendo su idioma más allá de la península.

### ***Danés***

El danés es una lengua muy importante debido a su influencia, ya que está estrechamente ligada a otras lenguas como el noruego y el sueco.

El idioma danés es hablado por aproximadamente 5 millones de personas. Su alfabeto se conforma de 29 letras, 26 de ellas tomadas del latín y 3 adicionales (Æ, Å y Ø). Su historia se remonta alrededor del año 1200, fecha en la que se encontraron los primeros textos de leyes en este idioma. Sin embargo, en Dinamarca se utilizaba con mayor frecuencia el latín para la redacción de textos administrativos. En el siglo XIV la reina Margaret I decidió sustituirlo por el danés y desde el siglo XV ha sido la lengua que se usa en la literatura<sup>59</sup>.

Actualmente el danés es hablado en el Reino de Dinamarca, aunque también es posible encontrar hablantes de danés en Alemania. En *Flensburg*, que es un área limítrofe con Dinamarca, viven 50,000 personas bilingües, aunque el alemán es el que más se utiliza.

Hasta mediados del siglo XVII el danés fue la lengua de las provincias *Skåne*, *Halland* y *Blenkinge*, que formarían Suecia en 1658. Después se cambió del danés al sueco como lengua escrita de estas regiones, pero restos del danés han persistido hasta ahora en dialectos de *Skåne* (Escania) como el *Skånska*.

En el periodo de unión de Dinamarca con Noruega (1389 a 1814) el danés se convirtió en una lengua muy importante en este último hasta que, en el siglo XX (cuando este se independiza del Reino de Suecia) surge un movimiento nacionalista y se estandarizan sus dos lenguas *bokmål* y *nynorsk*.

El idioma danés también se sigue hablando en las colonias danesas del Atlántico Norte (Groenlandia e islas Feroe) y por emigrantes de Estados Unidos y Argentina.

Los dialectos daneses se han mantenido pese a la estandarización, sin embargo, a los hablantes de danés en Copenhague se les dificulta entender los dialectos del este y sur de

---

<sup>59</sup> [König E. van der Auwera J, 1994]

Jutlandia. Los dialectos daneses son: Selandés (*Sjællandsk*, danés de Selandia), *Jysk* (danés de Jutlandia), *Bornholmsk* (danés de *Bornholm*) y *Fynsk* (danés de Fionia).

Hasta años recientes, muchas personas de la tercera edad seguían usando el viejo estándar, anterior a 1948, en el que los pronombres iniciaban con letras mayúsculas y se utilizaba la grafía “aa” en vez de la *å*. El estándar actual se llama *ringsdansk* (danés del reino) y su nombre se debe a que es aceptable en toda Dinamarca.

Los sociolectos<sup>60</sup> de Copenhague han influenciado fuertemente a las variedades regionales, esparciéndose rápidamente e incluso han llegado al norte de Jutlandia.

Existe un estándar para descripciones fonéticas llamado “Estándar avanzado de Copenhague”, que ha ganado cierta importancia. Este estándar es hablado principalmente por la gente culta de clase media nacida en la época de 1940 que habitó en los suburbios al norte de Copenhague. Esta variación sociolecta, considerada como variación de edad/grupo, ha sido muy estudiada.

El danés y el sueco en la Antigüedad no se podían diferenciar con facilidad. Fue hasta la época de los vikingos (de 800 a 1050 d.C.) que ambas lenguas se fueron separando.

Entre 1380 y 1814 el territorio de Noruega perteneció al reino de Dinamarca, lo cual explica la relación entre el noruego y el danés. En consecuencia, surgiría uno de los estándares escritos de Noruega, el *bokmål*, basado en el danés.

A pesar de que Dinamarca, Suecia y Noruega son países con lenguas muy parecidas, los habitantes de estos dos últimos pueden presentar dificultad para comunicarse con la gente de Dinamarca (en especial los suecos) debido a su fonética.

Fuera de Dinamarca, el danés se habla en las islas Feroe y Groenlandia, aunque también es enseñado en Alemania (*Flensburg*) e Islandia. En este último, el danés junto con el inglés son las lenguas que se enseñan de manera obligatoria en las escuelas.

---

<sup>60</sup> Sociolecto. Manera en la que hablan las personas que pertenecen a un mismo grupo sociocultural.

## **Sueco**

El sueco (*svenska*) es una lengua de Suecia con aproximadamente 9 millones de hablantes, sin embargo, no era oficial en Suecia hasta el año 2009<sup>61</sup>. Poco antes (en el año 2005) se presentó una propuesta ante el parlamento de Suecia para que el sueco fuera el idioma oficial del país, misma que fue rechazada<sup>62</sup>. El idioma sueco es la lengua oficial en las islas *Åland*, con aproximadamente 25,000 hablantes, y es utilizado también en zonas al poniente y sur de Finlandia con alrededor de 300,000 usuarios (donde es uno de los dos idiomas oficiales junto con el finés, pues este territorio perteneció al reino sueco hasta 1809).

Hasta la era vikinga el sueco y el danés eran casi indistinguibles. Fue hasta por el año 800 en el este de Escandinavia que se empezaron a separar.

La historia del idioma sueco se divide principalmente en dos periodos: el sueco antiguo de la época vikinga, que va desde el año 800 hasta 1526, y el sueco moderno de 1526 hasta la actualidad.

El sueco no ha sido influenciado por lenguas de los inmigrantes, pero tiene prestamos lingüísticos del alemán, francés y principalmente del inglés.

Como sucede en otros idiomas, el sueco también tiene diversos dialectos. Algunos reciben el nombre de dialectos antiguos, mismos que se siguen hablando en áreas rurales del norte de Suecia y en Finlandia. Otros dialectos son llamados urbanos y son muy parecidos al estándar central. También hay variantes locales, y como en la mayoría de los países, la variante de mayor estatus es la de la capital (en este caso la de Estocolmo). Otra variante importante es la que se encuentra al sur de Suecia (territorio que perteneció a Dinamarca).

## **Noruego**

El noruego cuenta con dos variantes o estándares escritos: el *bokmål* y el *nynorsk*. Estos se crearon a partir de la separación de Noruega del Reino de Dinamarca, ya que este último había impuesto al danés como la lengua de Noruega desde el año 1380 hasta 1814 en detrimento del noruego. Durante este lapso de tiempo surgieron varios dialectos noruegos

---

<sup>61</sup> <http://www.thelocal.se/20090701/20404>

<sup>62</sup> <http://www.svt.se/nyheter/inrikes/svenskan-blir-inte-officiellt-sprak>

que se utilizaban únicamente de forma oral. Al integrarse Noruega a Suecia empieza a haber un movimiento lingüista para retomar al idioma noruego con lo que se crean los estándares antes mencionados. El *bokmål* (lengua del libro) es hablado aproximadamente por el 85 a 90% de la población. Fue el lingüista Knud Knudsen (1812-1895) quien buscó transformar al danés integrando elementos del noruego hablado coloquialmente por las clases educadas que habitaron las áreas urbanas. El otro estándar es el llamado *nynorsk* (noruego nuevo) que fue creada por el filólogo y poeta Ivar Aasen (1813-1896) y se diseñó a partir de la unión de diferentes dialectos noruegos similar al nórdico antiguo. Inicialmente fue llamado noruego o *landmål* (la lengua del reino)<sup>63</sup>.

En 1885 el parlamento noruego decidió que tanto el dano-noruego como el noruego nuevo fueran considerados con igualdad en asuntos oficiales y en la educación.

Al dano-noruego en un principio se le llamó como *det almidelige Bogsprog* (la lengua literaria común) y después se le cambió la denominación a *riskmål* (la lengua del reino). Los términos actuales para los dos estándares se presentaron por primera vez en 1929.

En la actualidad, la variante *bokmål* es hablada por la mayoría de la población, mientras la variante *nynorsk* es menos utilizada en ámbitos militares, de alto grado académico y en publicaciones. Sin embargo, el *nynorsk* es más usado en áreas rurales y en algunos distritos del oeste de Noruega<sup>64</sup>.

Existen distintos dialectos del *bokmål/riskmål* entre estos, el más prestigiado es el que está basado en la forma de hablar de los habitantes de Oslo, capital de Noruega. Para el *nynorsk* no se tiene un estándar de pronunciación, es hablado con cualquier dialecto.

## ***Islandés***

El idioma islandés descende del nórdico antiguo y es el idioma que menos ha cambiado (en cuanto a morfología y sintaxis) de entre las lenguas escandinavas/nórdicas. Su

---

<sup>63</sup> [König E. van der Auwera J, 1994]

<sup>64</sup> <http://www.aasentunet.no/11+Offentleg+forvaltning.d25-SxdLK0y.ips?get=content>. Se puede consultar también en <https://www.ethnologue.com/language/nor>.

vocabulario se basa más en palabras nativas que en préstamos de otros idiomas y es la única lengua viva que aún utiliza la letra rúnica *P*; el idioma más cercano al islandés es el feroés.

En Islandia la gente estudia alemán, sueco y noruego, pero también se estudian obligatoriamente inglés y danés, sin embargo, la enseñanza de éste último se realiza de tal manera que sea más comprensible para los suecos y noruegos, por esta razón comúnmente es llamado *skandinaviska* (escandinavo) por los islandeses.

Los primeros asentamientos permanentes en Islandia aparecieron entre los siglos IX y X, siendo la mayoría de los colonos originarios de la actual Noruega. Se cree que estos pobladores hablaban distintos dialectos de nórdico antiguo pero que con el paso del tiempo se fueron homogeneizando hasta formar el islandés antiguo. Los lingüistas creen que aún existen rastros del nórdico occidental que han prevalecido en el islandés moderno dado que este es cercano al feroés y a los dialectos del suroeste de Noruega<sup>65</sup>.

La historia del islandés se divide en dos periodos: El periodo del islandés antiguo (antes de 1540) y el correspondiente al islandés moderno (después de 1540). Esta división se basa en la publicación de la primera traducción de la Biblia en islandés, en 1540.

En el islandés moderno existen algunas diferencias dialécticas entre pueblos, pero estas son menores en comparación de las demás lenguas germánicas y por lo tanto no hay estándares particulares o tipos de pronunciación definidos.

## 4.2.2 Germánico Occidental

### *Alemán*

El alemán es una lengua hablada en Europa Central determinada en gran parte por el estándar escrito que ha impuesto el gobierno alemán.

Existen diferentes tipos de pronunciación del alemán según su área geográfica, pero todos los alemanes pueden comunicarse a través de la gramática estándar, aunque en la vida cotidiana solo una pequeña fracción de los hablantes utiliza dicha forma, debido a que puede

---

<sup>65</sup> [König E. van der Auwera J, 1994]

ser impráctica y a que la población de diversas áreas está acostumbrada a sus propios sociolectos.

Los dialectos del alemán se clasifican en *Niederdeutsch* (bajo alemán), *Mitteldeutsch* (alemán central) y *Hochdeutsch* (alto alemán). El estándar moderno está basado principalmente en un dialecto del alemán central occidental (*Meißner Deutsch*).

La mayoría de los hablantes de alemán se encuentran en Alemania, Austria, Suiza, Luxemburgo y Liechtenstein, donde es una de las lenguas oficiales (es hablado como lengua materna por aproximadamente 100 millones de personas). También es utilizado en Bélgica, la región de Alsacia/Lorena (Francia), Tirol del Sur (Italia), Polonia, República Checa, Eslovaquia, Hungría, Croacia, Eslovenia y Rumanía entre otros países. Por último, cabe señalar que también se habla en América y Australia por emigrantes que llegaron durante los siglos XIX y XX, y que incluso llegó hasta Namibia (África) durante la época colonial<sup>66</sup>.

El alemán es uno de los idiomas más hablados como segunda lengua extranjera (colocándose por debajo del inglés) compitiendo con los idiomas chino, español, francés y ruso. Además, el número de personas que aprenden alemán se ha ido incrementando desde la unificación de Alemania, debido a que esta nación está entre la más influyentes de la Unión Europea desde entonces.

### ***Neerlandés***

La palabra en español neerlandés deriva de la palabra francesa *Néerlandais* que a su vez se deriva de la palabra *Nederlands* que es un adjetivo en este idioma que significa Tierra Baja o País Bajo.

Esta lengua está relacionada con el *Niederdeutsch* (bajo alemán) y es la lengua oficial de los Países Bajos, Bélgica, Surinam y las Antillas Holandesas. También es hablado como lengua materna en la región del bajo Rin de Alemania y en el norte de Francia<sup>67</sup>, de esta manera es la lengua materna de aproximadamente 23 millones de personas. También se habla

---

<sup>66</sup> Moreno-García J (1976) “La conferencia de Berlín (1884-1885)”, Historia 16

<sup>67</sup> [Willemyns R, 2002]

en Canadá (con aproximadamente 129,000 hablantes), Estados Unidos, Nueva Zelanda (con alrededor de 29,000 hablantes) e incluso Indonesia.

Esta es una lengua predominantemente germánica, lo que la hace más cercana al alemán que al inglés (grupo lingüístico germánico occidental) debido a que el inglés tiene influencias significativas del francés. A pesar de su similitud con el alemán, el sistema consonántico del neerlandés es más parecido al de las lenguas anglo-frisonas o escandinavas.

A este idioma en inglés se le designa como *Dutch*, que proviene de la antigua palabra germánica *Diet* cuyo significado es “pueblo”, ya que en esa época la gente culta hablaba latín. La palabra *Diet* también dio origen a la palabra en alemán *Deutsch* que significa alemán. En la Edad Media se le llamo *Diets*, *Dietsch*, *Duits* o *Duitsch* y en el renacimiento se cambió su denominación a *Nederduits* para diferenciarse de las lenguas vecinas alto alemán y bajo alemán (que como ya se mencionó, actualmente son llamadas *Hochdeutsch* y *Niederdeutsch*) que se apropiaron de la palabra *Duits* para designar al alemán de aquel periodo. El nombre *Nederlands* (neerlandés) es un nombre relativamente reciente y por esta razón no es popularmente reconocido en la mayoría de los países (en su respectivo idioma), siendo más populares los nombres holandés y flamenco para designarlo en los Países Bajos y Bélgica respectivamente<sup>68</sup>.

El neerlandés cuenta con dialectos que son mutuamente inteligibles entre sí, los cuales son *Hollands* (holandés), *Vlamms* (flamenco), limburgués y afrikáans.

El holandés, como su nombre lo indica, es un dialecto de las regiones de Holanda Septentrional (Holanda del Norte) y Holanda Meridional (Holanda del Sur) regiones pertenecientes a los Países Bajos. El flamenco se habla principalmente en la Región Flamenca (*Vlaams Gewest*), una de las tres regiones de las que se compone Bélgica. El limburgués es hablado en la provincia de Limburgo (en las ciudades de *Venlo*, *Mestreech*, *Hasselt*, *Genk*, *Mezeik* y *Voeren*) que pertenece a la Región Flamenca, además en Renania del Norte-Westfalia (en las ciudades de Aquisgrán y *Düsseldorf*) en Alemania. Finalmente, el afrikáans es hablado en Botsuana, Namibia y Sudáfrica (países del continente africano)<sup>69</sup>.

---

<sup>68</sup> [König E. van der Auwera J, 1994]

<sup>69</sup> <https://www.ethnologue.com/language/afn>

El neerlandés era hablado en la desembocadura de los ríos Rin, Escalda y Mosa en una zona que actualmente pertenece a los Países Bajos, Bélgica, Luxemburgo, norte de Francia y oeste de Alemania, que desde 1482 hasta 1581 fue llamada *Zeventien Provinciën* o las Diecisiete Provincias. Estas fueron gobernadas por el rey Carlos I de España (Carlos V del Sacro Imperio Romano Germánico) y posteriormente por su hijo el rey Felipe II de España. En 1555 el rey Carlos V abdicó a favor de su hijo Felipe para gobernar España y sus territorios, entre ellos las Diecisiete Provincias. Estas últimas eran tierras con una economía buena por sus actividades mercantiles por lo que el rey decidió incrementar los impuestos para seguir financiando el expansionismo español, también quería derrocar al calvinismo (una forma de protestantismo que se extendió por Francia y los Países Bajos) en esta región en favor de la iglesia católica. Los pobladores, cansados con estas decisiones empezaron a rebelarse en contra del reino español y en 1568 comienza la guerra de los Ochenta Años - inicialmente en la región sur del territorio-. El 5 de enero de 1579 fue firmado el *Tratado de Arrás*, en el cual los Países Bajos del Sur (Brabante, Limburgo, Luxemburgo, Artois, Henao, Flandes, Namur, Zutphen, Amberes y Malinas) juraron lealtad al rey Felipe II de España y reconocerían a Don Juan de Austria como su gobernante, además de la abolición de cualquier religión distinta de la católica lo que incluía al calvinismo. Como consecuencia, los protestantes liderados por Guillermo de Orange firmaron la *Unión de Utrecht* el 23 de enero de 1579, lo que dio origen a las Provincias Unidas o República de los Siete Países Bajos Unidos (Frisia, Groninga, Güeldres, Holanda, Overijssel, Utrech y Selandia). Lo que buscaba este acuerdo era el respeto a las tradiciones y creencias de la región, aunque seguirían siendo parte del reino español, sin embargo, el saqueo a la ciudad de Amberes por parte de los españoles por problemas económicos del reino y la persecución de Guillermo de Orange a quien se le consideró rebelde a la corona, fueron los principales motivos para que las Provincias Unidas (hoy en día Países Bajos) buscaran su independencia de España, la cual lograron el 26 de julio de 1581 con la firma del *Acta de abjuración* y el 24 de octubre de 1648 el reino de España reconoce su independencia mediante el *Tratado de Münster*.<sup>70</sup>

---

<sup>70</sup> Grattan T., Hawthorne J. (1899) *Holland: the history of the Netherlands*.

## **Inglés**

El inglés es el tercer idioma más hablado como lengua materna después del chino mandarín y el castellano. Muy probablemente la más importante en términos del alcance geográfico, porque es utilizada como *lingua franca* (permite la comunicación entre individuos que hablan diferentes lenguas). Es oficial y cooficial en 67 países y es hablado como lengua materna por aproximadamente 335 millones de personas y como la segunda por aproximadamente 505 millones de personas siendo así el idioma secundario más utilizado en todo el mundo después del chino mandarín.<sup>71</sup>

El inglés ha experimentado diversos cambios desde sus inicios (en el año 450 d.C.) hasta la actualidad. Este se clasifica en cuatro periodos: el primero comprende de los años 450 a 1150 el cual inicia con la llegada a las islas británicas de las tribus germanas anglos, sajones y jutos; en este periodo aparece el inglés antiguo. El segundo de 1150 a 1500 es la época del inglés medio, cuando empieza a formarse una lengua criolla debido a la conquista de Bretaña por parte de normandos que instauraron su idioma, el francés, el cual influyó fuertemente en el idioma inglés durante este periodo. El tercero que se sitúa entre los años 1500 a 1700, donde apareció el inglés moderno temprano o también llamado del renacimiento o isabelino, cuyas obras literarias más sobresalientes fueron escritas por *William Shakespeare*. Por último, se encuentra el inglés moderno, el cual inició en el año de 1700 y continúa hasta ahora.

### **Inglés Antiguo**

Los primeros pobladores de la Gran Bretaña fueron los celtas, quienes fueron invadidos por los romanos en el siglo I. Tras el declive del Imperio Romano, durante los siglos V y VI los anglos, sajones y jutos<sup>72</sup> (llegados de Schleswig, Alemania; Holstein, Alemania y Jutlandia, Dinamarca respectivamente) empezaron a invadir la isla con lo cual muchos britones huyeron hacia Gales, Cornwall y Escocia; algunos otros llegaron a Armórica (actualmente Bretaña, Francia), por lo que en ese territorio se habla bretón, que es una lengua

---

<sup>71</sup> Consultado en <https://www.ethnologue.com/language/eng> y [König E. van der Auwera J, 1994]

<sup>72</sup> Los jutos fueron los primeros pobladores de origen germánico en llegar a la isla desde Jutlandia en la actual Dinamarca. Sus líderes fueron los hermanos Hengist y Horsa quienes se aliaron con Vortigern el rey de los britones para combatir contra los pictos y escoceses entre los años 446 y 454. Este como agradecimiento les dio tierras en Kent.

celta, y hacia Galicia (en el actual España). Los únicos registros históricos que se tienen de las invasiones son los relatos de Bede “el Venerable” quien fue un teólogo e historiador anglosajón, sin embargo, este personaje confundió a los anglos y los sajones en sus narraciones. [Los anglos obtuvieron esa denominación porque su lugar de origen se llamaba *Engle* y a su lengua la nombraron *Englisc*, de esta se derivaría la palabra *English* (el idioma inglés)].<sup>73</sup>

Los vestigios más antiguos que se han encontrado del inglés son inscripciones rúnicas<sup>74</sup> anglo-sajonas que están fechadas en el siglo V.<sup>75</sup>

Durante mucho tiempo se desarrollaron dialectos del inglés, tales como el *northumbrian* en *Northumbria* (al norte del río Humber), el *mercian* en el reino de Mercia, el sajón occidental (*West Saxon*) en el reino de *Wessex* y el dialecto de *Kent* o *Kentish*. Durante los siglos VII y VIII la cultura y la lengua dominantes en Bretaña fueron las de *Northumbria*<sup>76</sup> hasta la invasión de los vikingos en el año 865 [liderada por Ivar Ragnarsson o Ívarr inn Beinlaus (Ivar el Deshuesado)] quienes dominaron gran parte de la isla de la Gran Bretaña a excepción del reino de *Wessex* quienes no solo se resistieron a los invasores, sino que recuperaron y conquistaron los demás territorios de la isla en 927 con lo que los reyes de *Wessex* serían también los reyes de Inglaterra e impondrían al sajón occidental como la lengua oficial del reino<sup>77</sup>. A la escritura de este periodo se le conoce como inglés antiguo, anglo-sajón o *Ænglisc*, cuyo alfabeto en un principio fue el rúnico *Futhorc* derivado de las lenguas escandinavas. Fue hasta tiempo después que los misioneros cristianos llevaron a Irlanda la escritura latina, por lo tanto su alfabeto consistía de caracteres latinos (sin la W que sería creada después en la época del inglés Medio), además se incluyeron las letras *Ææ* y *&* (ampersand que sustituía a la conjunción *and*), del anterior alfabeto se conservaron las letras rúnicas *Ðð* (*eth*) y *Pþ* (*thorn*) que corresponden al sonido de las letras *th* en el inglés moderno que se utilizaron de forma indistinta, y *ƿ* (*wynn* o *wen*) que sustituía al uso de la doble *v* (*vv*)

---

<sup>73</sup> <https://en.oxforddictionaries.com/definition/english>

<sup>74</sup> Las inscripciones rúnicas se grabaron en distintos objetos hechos de metal, piedra, madera e incluso hueso.

<sup>75</sup> [Orrin W. Robinson, 1992]

<sup>76</sup> <https://global.britannica.com/topic/English-language/Historical-background#ref603923>

<sup>77</sup> <https://global.britannica.com/place/Wessex-historical-kingdom>

o doble u (uu) por cuestiones de espacio. También se utilizaban las letras ȝ (yogh) que equivale a la letra “y” en el inglés moderno.

El vocabulario del inglés antiguo estaba compuesto por palabras del propio anglosajón, así como de algunas palabras de origen escandinavo y latín.

Uno de los escritos más reconocidos de esta lengua es el poema épico anglosajón *Beowulf*. El poema está dividido en dos partes: en la primera el joven gauta *Beowulf* ayuda a los daneses (jutos) para liberarlos de *Grendel*, un *jötunn*<sup>78</sup> que había invadido al reino de *Heorot* (tierra de los daneses) a quienes atemorizó devorándolos sin que nadie pudiera derrotarlo, ya que ningún arma le causaba daño. Entonces, *Beowulf* derrota a *Grendel* en una batalla a mano limpia para después pelear y derrotar a la madre de *Grendel*, quien era aún más poderosa. Victorioso *Beowulf* regresa a su pueblo en *Götaland* y después se vuelve el rey de los gautas. En la segunda parte *Beowulf* pelea, cincuenta años después, contra un dragón que atacó a *Götaland*, ya que un ladrón de este pueblo se habían robado su oro. Para poder atacar al dragón, *Beowulf* dirigió a un grupo de guerreros que huyeron ante la presencia del dragón, quedando sólo *Wiglaf* a su lado, tras su batalla victoriosa, *Beowulf* resulta herido de muerte, así que agonizando le concede el trono y el oro del dragón a *Wiglaf*.

### Inglés Medio

En 1066 los normandos conquistaron la Gran Bretaña, instalando así su cultura y su lengua, el francés antiguo<sup>79</sup>, con lo cual el idioma inglés volvió a tener cambios con la adición de más vocablos de origen francés. También cambió la forma para pluralizar los sustantivos, mientras que originalmente en el inglés antiguo los sustantivos en plural, por lo general, tenían terminación “-n” o “-en” (como en muchos de los casos del alemán actual) se cambió por la forma francesa, agregando “-s” o “-es” al final de los sustantivos. Solo unos cuantos sustantivos mantuvieron su forma plural original, algunos ejemplos son las palabras *men*, *oxen*, *feet*, *teeth*, *mice*, *children*, etc. El francés también afectó a las palabras con “cw” que se cambiaron por “qu” (por ejemplo, la palabra *cween* se convirtió a *queen*).

<sup>78</sup> En la mitología nórdica los *jötunn* eran gigantes que se encontraban principalmente en el *Jötunheimr*, uno de los nueve reinos de esta mitología, estos seres eran la contraparte de los dioses nórdicos.

<sup>79</sup> Los normandos hablaban normando, una de las lenguas consideradas como francés antiguo.

No fue hasta el siglo XIV que el inglés<sup>80</sup> fuera instaurado como lengua oficial nuevamente en la Gran Bretaña, cuando en 1399 por primera vez asciende un rey cuya lengua materna es el inglés y no el francés, dicho rey fue Henry IV.

Al final del siglo XIV el inglés medio fue el dialecto estándar de Londres. Geoffrey Chaucer, el poeta inglés más importante de la Edad Media y autor de *Los cuentos de Canterbury*, escribió sus obras en esta lengua a pesar de que la mayoría de sus contemporáneos ingleses preferían escribir en francés o latín debido a que se les atribuía mayor estatus. En la época del inglés antiguo se utilizaba el dígrafo “uu” que fue reemplazada por la letra rúnica wynn “ƿ ƿ” por los escribanos. En el año 1300 (época del inglés medio) se retomaría el uso del dígrafo “uu” pues la letra wynn era confundida con la letra “p” del alfabeto latino. Finalmente, este dígrafo se sustituyó por la letra “w” actual<sup>81</sup>.

### **Inglés Isabelino**

El inglés isabelino o inglés moderno temprano inició gracias al Renacimiento, el cual fue un movimiento cultural que se produjo en Europa Occidental en los siglos XV y XVI. Este trajo cambios favorables a las ciencias y artes debido a que se retomaron enfoques griegos y romanos de cómo ver al mundo. Debido a esto el latín, el griego y el francés fueron considerados como lenguas cultas y se produjo un gran entusiasmo sobre las lenguas “clásicas”. Con esto el idioma inglés entra a un nuevo periodo, en el que se tomaron muchos préstamos de palabras en latín. Estas palabras provienen directamente del latín o son palabras del latín de origen griego. Algunos ejemplos de palabras tomadas sin modificaciones son *genius, focus, nausea, paralysis*. Otras palabras provenientes del latín se alteraron como las palabras *anonymous, technique, biography*. También se introdujeron palabras con terminaciones o sufijos griegos “-ize” e “-ism” (como *actualize, terrorize, federalism, nationalism*, entre otras). Además, se adquirieron adjetivos inexistentes en esos momentos, algunos ejemplos son las palabras *marine* y *pedestrian*. Otros adjetivos se introdujeron como sinónimos de otros ya existentes como *masculine* y *feminine* como sinónimos de *manly* y *womanly* respectivamente o *paternal* para *fatherly*. Además de palabras en latín se

---

<sup>80</sup> El inglés nunca se dejó de hablar en la isla e incluso había seis dialectos de inglés antiguo que son: South Eastern, South Western, East Midland, West Midland, Northern y Lowland Scots. Este último evolucionó a inglés medio y después al escocés.

<sup>81</sup> [Freeborn D, 1992]

introdujeron frases del francés como *soi-sant* (supuestamente), *vis-à-vis* (contra), *sang-froid* (sangre fría) y palabras como *crépe*, *étiquette*, entre otras. En ocasiones se llegaron a tomar un número excesivo de términos tomados del latín por algunos estudiosos creando términos pedantes y hasta oscuros de los cuales muchos no sobrevivieron, algunos ejemplos de estas palabras son *suppeditate*, *illecebrous*, *fatigate*, etc. Solo algunas de estas palabras sobrevivieron y son usadas en la actualidad de manera cotidiana unos ejemplos son: *encyclopaedia*, *affability*, *mundane*, *extol*, etc. Algunos otros escritores sintieron que se estaba abusando del uso de palabras extranjeras y trataron de re introducir palabras viejas del inglés que fueron reemplazadas como el caso de *gleeman* en vez de *musician* o *inwit* en vez de *conscience* y crearon palabras nuevas con raíces germanas como *yeartide* sustituyendo a *anniversary*, *starlore* por *astronomy* o *speechcraft* por *grammar*. Pero la mayoría quedaron en desuso rápidamente.<sup>82</sup>

No solo las lenguas “cultas” influenciaron al idioma inglés, también intervinieron otro tipo de fenómenos como la masificación de la lectura y la escritura (que en la antigüedad eran privilegios que gozaban solamente unas cuantas personas como los gobernadores, clérigos y literatos), gracias a la imprenta que fue inventada en 1450 por Johann Gutenberg e introducida a Inglaterra en 1476 por William Caxton. De esta forma se imprimieron libros que en su mayoría eran cuentos míticos, poemas, gramáticas e historias populares, entre otros. La imprenta permitía realizar muchas copias de varios libros en poco tiempo, haciéndolos más asequibles para un mayor número de personas con lo cual los libros en inglés se empezaron a volver más populares que los escritos en latín que eran más costosos porque la mayoría de estos fueron hechos a mano.

A pesar de esto el uso de doblar la última letra de una palabra e incluso agregarle una “e” al final (como en *dog*, *dogg* o *dogge*) fue arbitrario por mucho tiempo. Uno de los cambios con el uso de la imprenta fue la eliminación de las letras rúnicas del inglés escrito, así la palabra “Pe” que se utilizó por siglos se empezó a escribir como “Ye” sin cambiar su pronunciación original (y posteriormente a *The*) debido a que la “Y” y la letra *P* eran muy

---

<sup>82</sup> <http://www.thehistoryofenglish.com> Crystal D. (1995) "The Cambridge Encyclopedia of the English Language"

parecidas. También se hizo una distinción entre las letras “U” y “V” que antes eran usadas indistintamente, de la misma manera con las letras “I” y “J”.

La biblia de rey Jacobo (*King James' Bible*), es quizás el trabajo más influyente en el desarrollo del idioma inglés, idiomáticamente hablando, e incluso tiene muchas frases que han pasado al español (Al este del Edén, dentro de la pureza todo es puro, lo que dios ha unido el hombre no lo puede separar, blanco como la nieve, etc, etc.). Esta biblia fue el primer texto en inglés publicado masivamente y por lo mismo tuvo (y sigue teniendo) una gran influencia en el desarrollo del idioma. Pete Seeger por ejemplo, tomó hace apenas 50 años tomó parte del libro de Eclesiastés 3:1-15 para la canción Turn! Turn! Turn! Entre ellas la frase “to everything, there is a season” (para todo, hay un momento adecuado), la cual ha pasado al idiomático actual desde entonces.

El escritor más prolífico, y que mayor influencia tuvo en el desarrollo del idioma inglés isabelino fue William Shakespeare, al cual el *Oxford English Dictionary* le atribuye haber introducido 3,000 palabras al lenguaje, aunque algunos estudiosos le atribuyen más de 17,000 incluidas palabras que están en desuso [Pressley 2015]. De acuerdo a esta misma fuente, el Isabelino puede considerarse como un dialecto del inglés moderno el cual tiene algunas anomalías preposicionales y de concordancia entre sujeto y verbo.

Pero no fue hasta por 1650 que se empezó a estandarizar realmente al idioma inglés. Este proceso se realizó de manera gradual siendo los nombres propios las palabras que más tardaron en estandarizarse dado que había muchas variantes, por ejemplo, existían al menos 20 formas de escribir “Shakespeare”<sup>83</sup> e incluso William Shakespeare los escribió de seis maneras distintas incluyendo dos en su testamento.

### **Inglés Moderno**

El inglés moderno inició en el siglo XVI. Una característica fue el cambio de las inflexiones o conjugaciones de los verbos en tiempo presente en tercera persona del singular, antes terminaban en “-th” ahora en “-s” por ejemplo “sche hath” cambió a “she has” y los verbos auxiliares se reemplazan por ejemplo “he is risen” cambió a “he has risen”. También

---

<sup>83</sup> <http://shakespeareauthorship.com/name1.html>

se cambiaron la mayoría de los pronombres personales y se eliminaron letras de su alfabeto dejando el *ampersand*, aunque este no se tiene oficialmente como letra del alfabeto actual.<sup>84</sup>

El inglés moderno acumuló varios cambios debido a la Revolución Industrial; puesto que en ese periodo se estaban haciendo muchas invenciones y aplicaciones, las cuales no estaban registradas en los diccionarios, así que el inglés tuvo que tomar prestamos lingüísticos de ellas en otros idiomas (*magnetic = magnetique* del francés) otras palabras se inventaron de raíces griegas y latinas (*vaccine, electron, protein, etc.*). Otras más fueron inventadas para objetos de reciente invención: *railway, locomotive, factory, etc.*

Un factor muy importante para la difusión del inglés fue el expansionismo colonial que tuvo Inglaterra en el siglo XVI, tanto en América como en Asia y Oceanía. Como resultado, se importaron muchas palabras asociadas a los productos, animales y plantas que venían de sus colonias [de Australia (*kangaroo, boomerang*) y de la India (*pyjamas, bungalow, jungle, shampoo, candy, etc.*)] o de las colonias españolas (*chocolate, mosquito, coyote, tomato, etc.*).

Gracias al expansionismo inglés, su idioma llegó al continente americano donde se empezó a usar en sus colonias, principalmente Canadá y Estados Unidos. Con la independencia de este último empezarían a llegar migrantes de varias partes de Europa, principalmente Holanda, Irlanda y Alemania quienes influenciarían al inglés tanto en su vocabulario como en la gramática.

---

<sup>84</sup> <http://www.thehistoryofenglish.com> Crystal D. (1995) "The Cambridge Encyclopedia of the English Language"

## 5. Marco de referencia

### 5.1 Filología comparada y familias de lenguas

Como ya se mencionó, la filología es la ciencia que se encarga de estudiar la lengua y la literatura de un pueblo a través de sus textos escritos. Esta ciencia surgió en el siglo XVIII cuando Sir William Jones descubrió la relación que existe entre el sánscrito y algunas lenguas europeas como el latín, griego y alemán. En 1885 se publicó el libro *Franz Bopp Comparative Grammar* donde se empezaron a tratar los conceptos de filología comparada y familias lingüísticas, definidas como grupos de lenguas que surgieron a partir de otra llamada “protolengua”.

Los lingüistas han ideado diversos métodos para reconstruir lenguas antiguas. Todos tienen como punto de partida las lenguas actuales, buscan similitudes en cuanto a gramática, sintaxis, vocabulario y vocalización, posteriormente con esta información van reconstruyendo los antepasados inmediatos de las lenguas estudiadas hasta alcanzar la lengua originaria. En la actualidad los estudios de las lenguas indoeuropeas se enfocan en las familias itálicas, celtas, germánicas, bálticas y eslavas.

En 1822, Jacob Grimm formuló una ley llamada *Lautverschiebung* (cambio de pronunciación) que explica, entre otras cosas, por qué las lenguas germánicas conservan sus consonantes fuertes a pesar de que la mayoría tienden a conservar las más suaves.

A partir de esta ley los lingüistas empezaron a trabajar en la reconstrucción de un protolenguaje que llamaron indoeuropeo, basándose en la idea de que grupos de consonantes van desplazando a otros a través del tiempo de manera predecible y regular.

Los descendientes de la lengua protoindoeuropea son las lenguas más esparcidas en el mundo y con el mayor número de hablantes, esta familia ha sido estudiada durante más de doscientos años por varios investigadores que han ido reconstruyendo el vocabulario y sintaxis de la lengua protoindoeuropea, además han tratado de descifrar los caminos que tomó esta protolengua que se cree surgió en la región del Cáucaso, en el actual Azerbaiyán.

En la siguiente imagen se muestran algunas de las ramas de la familia de lenguajes indo-europeos donde, al final de cada rama se muestran las lenguas analizadas a lo largo de este trabajo.

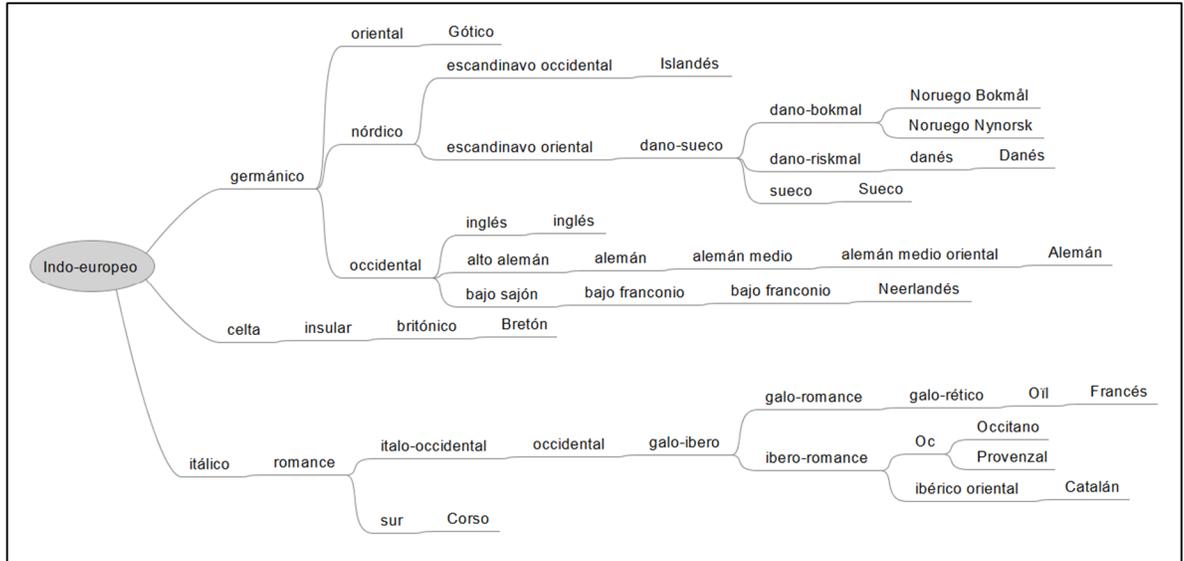


Figura 7. Parte de la familia indo-europea.

## 5.2 Modelo de coloración de gráficas suaves

### 5.2.1. Definición de coloración de gráficas suaves [Lara 2015]

En el problema de coloración de gráficas se tiene una gráfica conexa (una gráfica en la cual, dados cualquier par de vértices existe un camino entre ellos) no dirigida y se desea colorear todos sus vértices con el mínimo número de colores, evitando colorear dos vértices adyacentes con el mismo color.

La coloración de gráficas suaves CGS es un caso especial del problema de coloración de gráficas en donde además de tener una gráfica completa y no dirigida, también está ponderada (a los valores de las aristas en este problema se les llama penalizaciones), sin embargo, en este tipo de problemas el objetivo es distinto, aquí se busca encontrar la coloración (configuración) que minimice un valor asociado a la gráfica, llamado dureza. Este valor se calcula sumando las penalizaciones de todas las aristas cuyos vértices tengan el mismo color asociado. El problema se puede definir de la siguiente manera:

Sea  $G$  una gráfica completa, es decir, un conjunto de los vértices  $V$  en  $G$  y todas sus posibles aristas  $E (i, j)$ .

$$G = (V, E); |V|=n; |E|=n(n-1)/2$$

Definimos a la penalización de cada arista  $(i, j)$  denotado por  $p_{ij}$  tal que:

$$p_{ij} \geq 0, \forall (i, j) \in E,$$

Una función  $C^k$ -coloración de los vértices de  $G$  está definida como:

$$C^k: 1, 2, \dots, n \rightarrow 1, 2, \dots, k$$

Donde  $k$  es el número total de colores  $1 < k < n$  que identifica  $C(i)$  como el color del vértice  $i$ . Para una coloración  $C^k$  en una gráfica  $G$  la función de dureza de una coloración  $C^k$  está dada por:

$$H(C^k) = \sum_{(i,j) \in E, C^k(i)=C^k(j)} p_{ij}$$

El objetivo de la CGS es encontrar  $C_{op}^k$  de  $G$  que minimiza la dureza  $H(C_{op}^k)$ .

### 5.2.2. Solidez y resiliencia de una coloración

Dada una coloración con  $k$  colores en una gráfica completa con  $n$  vértices, el número promedio  $m$  de vértices pintados con el mismo color es  $m = n/k$  y el número de aristas compartidas por  $m$  vértices son las combinaciones  $C(m, 2) = m(m-1)/2$  por lo que el número promedio de penalizaciones incluidas en la función de dureza debe ser proporcional al número promedio de vértices pintados con el mismo color multiplicado por el número de colores. Definimos la función de solidez de una coloración como la dureza de un gráfico dividido por el número medio de aristas que contribuyen a la dureza:

$$s(C_{op}^k) = \frac{H(C_{op}^k)}{km(m-1)/2} = \frac{2H(C_{op}^k)}{k \frac{n}{k} \left( \frac{n}{k} - 1 \right)} = \frac{2kH(C_{op}^k)}{n(n-k)} \quad (1)$$

La resiliencia de una coloración  $C^k$  se define como el porcentaje en que disminuye la solidez de una coloración con  $k - 1$  colores con respecto a la realizada con  $k$  de ellos y se puede expresar como sigue:

$$R(C_{op}^k) = \frac{s(C_{op}^{k-1}) - s(C_{op}^k)}{s(C_{op}^k)} \quad (2)$$

Si resulta que, al añadir un color adicional a la coloración de la gráfica, hay un aumento significativo en su resiliencia, eso significa que hemos encontrado un número adecuado de clases para agrupar.

Finalmente, si obtenemos la resiliencia para todas las coloraciones consecutivas desde 1 hasta  $n$ , y si pensamos en cada color como una clase para agrupar a los vértices (objetos), los valores más grandes mostrarán las mejores opciones de número de clases que se pueden utilizar para agrupar el conjunto.

### 5.2.3. Distancia entre lenguas

El análisis de frecuencias de letras sirve para saber qué tan similares son dos lenguas. Una muy buena métrica que hace las diferencias más notorias es la distancia euclidiana cuadrada:

$$d_{i,j} = \sum_{k=1}^n (x_{ki} - x_{kj})^2$$

Donde  $X_{ki}$  es la  $k$ -ésima letra en la lengua  $i$ .

Pues las menores se hacen más pequeñas y las mayores más grandes. De esta manera, entre más chica sea la diferencia entre un par de lenguas, estas tienen mayor parecido.

## 5.3 Método de solución binario CGS [Lara 2015]

Dada una gráfica no dirigida  $G = (V, E)$  con  $|V| = n$  y  $|E| = n(n - 1)/2$  se desea obtener una coloración  $C_{op}^k$  usando  $k$  colores. Para resolver este problema definimos el modelo de programación binaria siguiente:

$$\min z = \sum_{(i,j) \in E} p_{ij} y_{ij}$$

sujeto a:

$$\begin{aligned} \sum_{l=1}^k x_{il} &= 1 \quad \forall i \in \{1, \dots, n\} \\ x_{il} + x_{jl} - 1 &\geq y_{ij} \quad \forall (i,j) \in E \quad \forall l \in \{1, \dots, k\} \\ x_{il} &\in \{0, 1\} \quad \forall i \in E \quad \forall l \in \{1, \dots, k\} \\ y_{ij} &\in \{0, 1\} \quad \forall (i,j) \in E \end{aligned} \quad (3)$$

donde:

$$x_{il} = \begin{cases} 1 & \text{si } C(i) = l \quad \forall i \in \{1, \dots, n\} \quad \forall l \in \{1, \dots, k\} \\ 0 & \text{en otro caso} \end{cases}$$

$$y_{ij} = \begin{cases} 1 & \text{si para alguna } l \in \{1, \dots, k\} \text{ se cumple que } x_{il} = x_{jl} = 1 \\ 0 & \text{en otro caso} \end{cases}$$

Donde el conjunto de restricciones del sistema (3) garantiza que cada vértice de la gráfica tenga asociado solamente un color  $l \in \{1, \dots, k\}$ ; la función  $y_{ij}$  vale 1 si los vértices  $i$  y  $j$  tienen asignado el mismo color y cero en otro caso. En el primer caso la penalización  $p(i, j)$  se toma en cuenta para el cálculo del valor de la función objetivo  $z$  (valor de dureza). Al resolver el problema de programación binaria además de encontrar el valor de dureza mínimo también se encuentra la coloración óptima, esto es:  $z^* = H(C_{op}^k)$ .

El sistema (3) tiene  $nk$  variables  $x_{il}$  y  $n(n-1)/2$  variables  $y_{ij}$  por lo que el número total de variables binarias en el modelo 1 es  $nk + n(n-1)/2$ . Por otro lado el número de ecuaciones es igual a  $n$  y el número de desigualdades es igual a  $kn(n-1)/2$ , por lo tanto, el número total de restricciones es igual a:  $n + kn(n-1)/2$ .

## 6. Estado del Arte

### 6.1 Modelo Filológico para lenguas ibéricas

En el trabajo “Modelo Filológico para lenguas ibéricas” de Pedro Lara [Lara 2015] se compararon y clasificaron diez de las lenguas más representativas de la Península Ibérica mediante una comparación estadística. Con esta se pudieron conocer cuales lenguas son más cercanas entre sí y el número de grupos en que se pueden clasificar.

En la Península Ibérica se encuentran distribuidas distintas culturas, cada una con su respectiva lengua. Estas lenguas son muy similares pero independientes entre sí, debido al aislamiento parcial de la Península Ibérica con el resto del continente europeo gracias a la frontera natural que forman los montes Pirineos.

Una lengua y su cultura son muy íntimas entre sí, esto se debe a que una cultura va forjando a una lengua con el paso de las décadas evolucionando juntas. Por ejemplo, el vocabulario que incluye palabras de alimentos, festividades, pueblos, animales de una región, etc. Ambas van definiendo la forma de pensar de los individuos de una sociedad y estos individuos pueden a su vez llegar a modificar su lengua y su cultura por distintas circunstancias, tales como la influencia de otras culturas y sus respectivas lenguas. Así que para tener una mejor comprensión de una cultura es necesario comprender su lengua o lenguas asociadas.

España es un país formado por varias regiones, cada una de las cuales tiene su propia lengua, cultura, política, economía y comportamiento social. Debido a esto los españoles se sienten más identificados con la región donde viven que con el país al que únicamente ven como un recaudador de impuestos. Un ejemplo claro es el País Vasco, que es considerada, una región autónoma con su propia lengua desde la constitución de 1978, sin embargo, desde la época de la dictadura de Francisco Franco su pueblo ha buscado la independencia de España. Otro ejemplo es el País Valenciano<sup>85</sup> que también es una región autónoma.

---

<sup>85</sup> <http://web.archive.org/web/20111005045527/http://www.cortsvalecnianes.es/descarga/destacados/ESTATU>

En contraste, Portugal es el país europeo más homogéneo en cuanto a su cultura y lengua, la única división del país es geográfica ya que está dividido en dos por el río Tajo quedando una zona norte y una sur. Portugal no es muy distante culturalmente de España debido a que ambos países forman parte de la Península Ibérica, misma que se encuentra aislada del resto de Europa por la cordillera de los Pirineos<sup>86</sup>.

Todas las lenguas de la Península Ibérica son lenguas hermanas que comparten la misma raíz, el latín, con excepción del euskera o vasco la cual es una lengua celta.

## 6.2 Reconocimiento de literatura inglesa

En el trabajo [Lara, 2015] se utilizó el modelo de coloración de gráficas suaves para el reconocimiento de textos de literatura inglesa y clasificarlos por épocas. En este trabajo se sigue el modelo de coloración de gráficas suaves utilizando como métrica la distancia euclidiana, la cual ya se ha mostrado ser la que arroja los mejores resultados sobre las distancias *Manhattan*, Euclidiana y Producto Punto.

Se utilizaron distintos textos para cada uno de los estadios del inglés (inglés antiguo, inglés medio, inglés isabelino e inglés moderno). Para el inglés antiguo se utilizaron dos de las obras más representativas de este periodo que son: *Beowulf* (Bw) y *La Crónica Anglosajona* (AC). En el caso del inglés medio se usaron: *Los Cuentos de Canterbury* (CT) y *Sir Gawain y el caballero verde* (SG). Las obras utilizadas para el inglés isabelino fueron: *Romeo y Julieta* (RJ), *Hamlet* (Hm), *Othello* (Ot), *Macbeth* (Mc) y *Volpone* (VI). Finalmente,

---

TO.pdf

<sup>86</sup> Durante siglos, en la península se asentaron distintos reinos, los cuales se fueron uniendo a partir de la fusión de los reinos de Castilla y Aragón con la boda de los herederos al trono de estos (Isabel I y Fernando II, apodados los Reyes Católicos). Pronto el reino español se empezó a expandir más allá de sus límites geográficos con el descubrimiento de nuevas rutas marinas, obteniendo cada vez más recursos y poder militar. En el año de 1580 tras la muerte de Enrique I de Portugal, uno de sus sobrinos, Antonio el Prior de Crato, se autoproclama como el nuevo rey de Portugal con el apoyo del pueblo, sin embargo, Felipe II de España al ser hijo de Isabel I de Portugal, reclama este reino, al ser rey de España y tener el apoyo del clero y la nobleza, manda un ejército liderado por Fernando Álvarez de Toledo y Pimentel, III duque de Alba para vencer a las tropas portuguesas. El rey Felipe II de España, tras la victoria de la Batalla de Alcántara es nombrado como Felipe I rey de Portugal, con lo cual se unificaría por primera vez la Península Ibérica, esta unión duraría 80 años debido a cambios de políticas hechas por Felipe IV quien aumentó los impuestos debido a los gastos que le generaban los distintos conflictos con los que tuvo que lidiar en el año 1640. En ese mismo año la nobleza portuguesa conspiró en contra de Felipe IV y nombró rey a Juan IV de Portugal, a partir de este hecho se inició un conflicto que duro 28 años (entre los reyes de España y Portugal) el cual terminó el 18 de febrero de 1668 con la firma del tratado de Lisboa en el cual se reconoce la independencia de Portugal. Fuente: Birmingham D. (2003) "A Concise History of Portugal"

para el inglés moderno se emplearon las siguientes: *Los viajes de Gulliver* (GT), *Frankenstein* (Fr), *Orgullo y Prejuicio* (PP), *La isla del doctor Moreau* (DM) y *Oliver Twist* (OT). Debido a que las obras en inglés antiguo y medio eran en promedio muy cortas, se utilizaron solamente dos de cada periodo ya que, las obras de unos pocos párrafos no permiten hacer una buena clasificación. Mientras que en el inglés isabelino y moderno existen muchas obras impresas debido a la invención de la imprenta a mediados del siglo XV, poco antes del periodo del inglés isabelino.

La selección de los idiomas se realizó mediante los siguientes criterios:

**Eng8.** Se seleccionaron 8 textos, de tal manera que se tienen dos textos correspondientes para cada periodo: Bw y Ac para el inglés antiguo, CI y SG para el inglés medio, RJ y Hm para el inglés isabelino y finalmente GT y Fr para el inglés moderno.

**Eng9.** Se escogieron nueve textos para tres clasificaciones: Bw, CT y SG para el inglés antiguo/medio; VI, Ot y Mc para el isabelino y PP, DM y OT para el inglés moderno.

Con los datos obtenidos de las instancias Eng8 se puede observar que el mayor valor de resiliencia se obtiene con la clasificación con seis colores y le siguen las de tres y dos colores. Sin embargo, se esperaba que el mayor valor de resiliencia se alcanzara con la de cuatro colores correspondiente al trabajo de los lingüistas, no obstante, con estos datos se puede deducir información adicional. Con el ordenamiento con tres colores, se puede observar que los textos en inglés antiguo e inglés medio se agruparon como se esperaba, mientras que los textos que históricamente corresponden al inglés isabelino e inglés moderno fueron puestos en un grupo, el del inglés moderno<sup>87</sup>. Para el caso con dos colores, pone juntos al inglés antiguo con el inglés medio y a parte al inglés isabelino con el inglés moderno. Para el caso de las instancias de Eng9 se esperaba que el número de colores con mayor valor de resiliencia fuera tres, sin embargo, el modelo indicó que la mejor clasificación fue con dos colores. De nuevo el modelo puso los textos correspondientes al inglés antiguo e inglés medio en un mismo grupo mientras que colocó a los textos en inglés isabelino e inglés moderno en otro grupo, ratificando el parecido de los textos dentro de algún grupo. La correspondiente a la de cuatro colores nos muestra los textos acomodados justamente por cada etapa del inglés.

---

<sup>87</sup> El inglés isabelino también es considerado como inglés moderno temprano.

Finalmente, la concerniente a seis colores asignó a los tres libros que corresponden lingüísticamente al inglés antiguo/medio a un grupo cada uno; estos pertenecientes al inglés moderno formaron un solo grupo como se esperaba; los restantes corresponden al inglés isabelino formaron dos grupos uno de ellos formado por las obras *Othello* y *Macbeth* ambas obras de *William Shakespeare*.

En conclusión, este trabajo muestra que, pese a no obtener los valores mayores de resiliencia para los números de colores esperados, el modelo ofrece muchos resultados interesantes, la mayoría de estos coincide con lo dicho por los lingüistas.

### **6.3 Un sistema Clasificador Utilizando Coloración de Graficas Suaves**

En el trabajo de Jorge Flores (Un sistema Clasificador Utilizando Coloración de Graficas Suaves) [Flores, 2016] se hizo uso de un sistema clasificador utilizando coloración de gráficas suaves, el cual es un sistema agrupador que no requiere de la intervención humana para realizar las clasificaciones. Para generar la matriz de distancias se requirió el seguimiento de las siguientes fases: 1) Limpieza de los datos faltantes: las bases de datos a utilizar se depuraron quitando las instancias a las que es faltara por lo menos un campo, 2) Ponderación de las columnas alfanuméricas: el sistema solo admite datos numéricos, por lo que los datos alfanuméricos fueron sustituidos por datos numéricos tomando en cuenta algunos criterios y 3) Normalización de datos: algunas columnas contaban con datos cuyas escalas diferían, pero el modelo considera que todas las columnas influyen de igual manera para la clasificación, por lo que se realizó una normalización de las bases de datos.

El modelo fue probado con cinco bases de datos distintas tomadas del Repositorio de Aprendizaje Automático de la Universidad de Irvine California. Estas bases de datos son “Hepatitis”, “Wine”, “Iris”, “Car Evaluation” y “Stone Flakes”.

La base de datos “Hepatitis” contiene los datos del tratamiento y síntomas de 155 pacientes de hepatitis. Cuenta con 19 características, entre las cuales se encuentran datos del paciente, tratamientos, síntomas, características del hígado, producción de sustancias en el cuerpo, entre otras. Esta base de datos está dividida en 2 clases que son “vivo” (con 123 instancias) y “muerto” (con 32 instancias). Los resultados muestran que la coloración con

mayor valor de resiliencia es para la de 2 colores que concuerda con el número de clases que ya se sabía a priori.

La base de datos “Wine” muestra los resultados del análisis químico de 178 vinos de una misma región, esta base de datos toma en cuenta 13 características que son básicamente las cantidades de distintas sustancias y el grado de alcohol en los vinos. Además, los datos están distribuidos en tres clases: vino de mesa (con 59 instancias), vino de crianza (con 71 instancias) y vino de reserva (con 48 instancias). Para esta base de datos el modelo sugiere que las instancias se agrupan en 3 clases, ya que el número de colores con mayor valor de resiliencia es 3, que coincide con los tipos de vinos. Por otro lado, se fueron tomando cierto número de instancias para medir, tanto tiempo de ejecución como porcentaje de eficiencia, con lo que con hasta 15 instancias se obtiene un 100% de eficiencia y tiempos de alrededor de los 6 segundos; para 60 instancias se obtuvo una eficiencia de 96.6%, tardando aproximadamente 60 segundos y para toda la base de datos (178 instancias) se obtuvo una eficiencia de 93.25% tardando 471 segundos

La base de datos “Iris” es probablemente la más reconocida en la literatura de reconocimiento de patrones. Los atributos de esta base de datos son: longitud de sépalo en cm, ancho de sépalo en cm, longitud de pétalo en cm y ancho de pétalo en cm. Los tipos de plantas se distribuyen entre iris setosa, iris versicolor e iris virginica.

El modelo muestra que el número de colores con mayor valor de resiliencia es 2 y el segundo color de mayor valor de resiliencia es 3. Estos resultados son congruentes con la información que ya se conocía como el hecho que existieran 3 tipos de plantas en la base de datos y que una de las clases se podía separar linealmente de las otras dos, por lo que un color correspondería a iris setosa mientras que el otro sería para iris versicolor y virginica. También se analizaron los tiempos de ejecución y porcentaje de eficiencia, que para 150 instancias se obtienen tiempos de ejecución de alrededor de 6 minutos con porcentajes de eficiencia entre 95% y 96%.

Para la base de datos “Car Evaluation” fueron evaluados 1728 modelos de automóviles según atributos subjetivos y técnicos con características como el precio, grado de mantenimiento, número de puertas, capacidad, tamaño de cajuela y grado de seguridad. Los

datos se encontraron distribuidos entre las siguientes categorías: inaceptable (con 1210 instancias), aceptable (con 384 instancias), bueno (con 69 instancias) y muy bueno (con 65 instancias).

Los resultados obtenidos por este modelo nos indican que el número de colores con mayor resiliencia es 2 seguido por 4 colores, esto quiere decir que el sistema obtiene mayores valores de resiliencia clasificando a los automóviles simplemente como buenos o malos según sus características, por otro lado, muestra que al clasificar a los automóviles en 4 clases se obtiene un valor de resiliencia alto que coincide con la clasificación original. Con respecto al análisis de tiempos de ejecución y porcentaje de eficiencia se obtuvo que se pueden alcanzar porcentajes de eficiencia de 100% con hasta 80 instancias con tiempos de ejecución de aproximadamente 100 segundos, mientras que con 200 instancias se obtuvieron porcentajes de 98.5% de eficiencia con tiempos de ejecución de aproximadamente 600 segundos.

La base de datos “Stone Flakes” fue publicada en 2014. Se compone de 79 objetos que sirvieron para crear puntas de flecha, lanzas o cuchillos en la Europa Occidental durante el periodo paleolítico. Con esta base de datos se busca encontrar relaciones entre el progreso tecnológico debido a su uso o el cambio en la técnica de creación de armas de la prehistoria. Esta base de datos cuenta con ocho atributos entre los cuales se encuentran las dimensiones de los objetos, frecuencia de uso, desgaste y ángulo de golpeo.

Esta base de datos no cuenta con clases definidas, sin embargo, los arqueólogos sugieren tres clasificaciones, las cuales toman en cuenta el tipo de homínido que las creó y su antigüedad: *Homo Ergaster* (600,000 y 300,000 años de antigüedad), técnica *Levallois* (200,000 años de antigüedad), Neandertales (130,000 y 60,000 años de antigüedad) y *Homo sapiens* (40,000 años de antigüedad).

El modelo sugiere que el mejor número de colores para clasificar es 2, aunque también 3 colores resultaron buenos para la clasificación. Estos resultados indican que para dos colores el modelo puede diferenciar las herramientas más antiguas de las más recientes y para el caso de tres colores puede diferenciar las herramientas del Bajo Paleolítico y *Levallois*, pero no puede diferenciar entre herramientas de los Neandertales y *Homo sapiens*.

## 7. Modelo Filológico propuesto

La distribución estadística de caracteres ASCII de textos es particular para cada idioma, sin embargo, puede mostrar similitudes al compararse entre idiomas que guarden cierto parecido. Por lo que se pueden comparar dos lenguas, entre más pequeñas sean las diferencias de las frecuencias de las letras en textos, más cercanas serán las lenguas y podrían estar emparentadas compartiendo las mismas raíces, es decir, podrían tener una o más proto lenguas en común. Este hecho se descubrió poco después de la invención de la imprenta en el siglo XV.

En este trabajo se buscó clasificar lenguas germánicas y francesas utilizando un modelo matemático que permite encontrar a las que más se parecen, por medio de la distribución estadística de caracteres ASCII en textos de cada una. Para esto, se obtienen las frecuencias de apariciones de las letras en cada idioma y estas se comparan por pares de idiomas restando las frecuencias y los resultados se eleva al cuadrado, esto se hace con la finalidad de tener valores positivos y para que las diferencias sean más notorias -los valores pequeños se vuelven más pequeños y los grandes aún más grandes-. Una vez que se tienen los valores para todos los caracteres estos se suman. El resultado nos indica que tan cercanos son dos idiomas, a este valor se le llama penalización. En el modelo se utilizan las penalizaciones entre cada par de lenguas para poder calcular un valor llamado dureza para las distintas coloraciones dado un número de colores.

Para encontrar la dureza mínima y sus coloraciones asociadas para cada número de colores, se utilizó un programa que realizó búsquedas locales, este algoritmo encuentra los valores mínimos en el 95% de las veces que se ejecuta.

Los datos que se obtienen son la dureza y su coloración óptima asociada (para cada número de colores), sin embargo, dados estos datos todavía no se puede observar cual es la mejor manera de clasificar, para saber esto se utilizan los valores obtenidos, con estos se

calculan las solideces de cada gráfica. La solidez es la dureza de una gráfica dividida por el número de aristas que contribuyen a esta.

A partir de las solideces se puede calcular otro valor llamado resiliencia. La resiliencia es el porcentaje en que disminuye la solidez de una coloración con  $k-1$  colores con respecto de otra con  $k$  de ellos.

Entre más alto sea el valor de resiliencia asociado a un número de colores más adecuado será el número para clasificar.

La métrica utilizada para obtener las distancias entre lenguas (penalizaciones de las aristas) es la distancia euclidiana cuadrática:

$$d_{i,j} = \sum_{k=1}^n (x_{ki} - x_{kj})^2$$

Donde  $X_{ki}$  es la  $k$ -ésima letra en la lengua  $i$ .

En otras palabras, se obtienen las diferencias de las frecuencias de cada carácter en cada par de lenguas, se elevan al cuadrado y se suman.

El hecho de elevar al cuadrado las diferencias tiene dos propósitos: el primero, es para no tener valores negativos que al momento de sumarlos den una distancia negativa y el segundo, es para que las diferencias pequeñas sean aún más pequeñas y las diferencias grandes sean más grandes.

Esta métrica se utilizó con éxito en el trabajo de clasificación de lenguas ibéricas [Lara 2015].

Las instancias que se utilizaron para clasificar fueron textos simples, i. e., textos sin formato, con codificación UTF-8, con una longitud entre 700 kB y 3300 kB para las lenguas

germánicas (alemán, neerlandés, inglés, gótico, danés, islandés, noruego *bokmål* y noruego *nynorsk*) y para las lenguas francesas (francés, vasco, catalán, bretón, occitano y corso).

Para obtener los valores óptimos de dureza y sus coloraciones asociadas se utilizó la búsqueda local:

1. Se crea una coloración al azar (cadena de longitud  $n$  que utiliza  $k$  números, donde  $k < n$ ) y se evalúa.
2. Se crea una copia de la coloración actual y se modifica seleccionando una posición indistinta dentro de la cadena y el número apuntado se modifica aleatoriamente.
3. Se evalúa la nueva coloración y se comparan ambos valores de dureza.
4. Se elige a la coloración cuyo valor de dureza sea el menor y se repiten el algoritmo desde el paso 2  $m$  veces.

## 7.1 Algoritmo de Solución

El algoritmo de solución cuenta de dos etapas: la primera es obtener la matriz de distancias para cada familia de lenguas y la segunda consiste en obtener las coloraciones y sus respectivas durezas óptimas (mínimas) para cada número de colores.

Algoritmo para hacer el conteo de caracteres para cada lengua.

Abrir listaLenguas.txt como (listaL)

Hacer

lengua = leelinea(listaL)

lenguas+=1

Abrir lengua como (len)

Hacer

cadena=leelinea(len)

cadena=minúsculas(cadena)

Para j=1 hasta 26

Para i=0 hasta longitud (cadena)-1

Si cadena[i]=j+96 entonces

arreglofrecuencias[lenguas,j]+=1

Fin Si

Fin Para

Fin Para

Hasta fin\_del\_archivo(len)

Hasta fin\_del\_archivo(listaL)

Una vez que se tienen los números de apariciones de cada letra en cada lengua, se obtienen las frecuencias de cada letra en cada idioma como se describe en el siguiente algoritmo:

Para i=0 hasta lenguas

suma=0

Para j=1 hasta 26

suma+=arreglofrecuencias(i,j)

Fin Para

Para j=1 hasta 26

arreglofrecuencias(i,j)= arreglofrecuencias(i,j)/suma

Fin Para

Fin Para

Por último, se obtiene la matriz de distancias, donde se obtuvieron distancias euclidianas cuadradas.

Para i=0 hasta lenguas

Para j=0 hasta lenguas

Para k<=1 hasta 26

matrizdistancias(i,j)+= (arreglofrecuencias(i,k)-  
arreglofrecuencias(j,k))^2

Fin Para

Fin Para

Fin Para

El siguiente es el algoritmo que corresponde a la segunda etapa:

Se crea el objeto “coloracion”

Se crea la matriz “colaux”

Se carga la matriz de distancias desde un archivo

coloración=generaColoracion //se crea una coloración aleatoriamente

coloración=dureza(coloración, vectorDistancia //se calcula la dureza de la coloración

Para ciclo=0 hasta 25

Hacer

color=numero\_aleatorio\*numeroColores+1

indice=numero\_aleatorio\*numeroLenguas

colorAuxiliar[indice]=color

mientras evaluaDureza(coloración, matrizDistancias)<evaluaDureza(colorAuxiliar,  
matrizDistancias)

coloracion=coloracionAuxiliar

coloracion=evaluaDureza(coloración, matrizDistancias)

Fin Para

Imprimir (coloracion)

Imprimir(dureza)

## 8. Instancias

### 8.1 Familia de lenguas francesas

El francés es el único idioma reconocido como oficial por el gobierno de Francia. Desde sus inicios fue la lengua de los gobernantes y de la elite francesa, por lo que ha tenido un prestigio muy alto en Francia y en el mundo. Antes de la Segunda Guerra Mundial, fue una lengua franca que facilitaba la comunicación entre países con distintos idiomas. Su situación era similar al inglés en la actualidad, se utilizaba para el comercio, la difusión científica, etcétera. Sin embargo, el francés no ha sido la única lengua que se habla en Francia, ha coexistido con otras lenguas y sus dialectos como el occitano, catalán, bretón, picardo, por mencionar algunos. Pero al igual que otros países europeos, el gobierno francés buscó crear un nacionalismo y promoverlo entre los habitantes franceses, esto con el fin de tener una nación unida y estable. Esto para evitar que distintas comunidades se llegaran a sentir ajenas a Francia y decidan independizarse como ha sido el caso de distintas comunidades de España, o el caso de Texas que se independizó de México. Para lograr este nacionalismo, el gobierno de Francia reconoció al francés como la única lengua oficial y cualquier otra lengua originaria de Francia sería, no solo no reconocida, sino que también desprestigiada y se le designaría con el sustantivo despectivo de “Patois”<sup>88</sup>, es decir, una lengua (y no un grupo de lenguas francesas) de gente inculta y hasta ruin, una lengua incompleta de gente bárbara. Si una persona buscaba conseguir un trabajo en el gobierno, en la educación, o algún otro buen empleo debía hablar francés como requisito principal.<sup>89</sup> Además, hubo decretos gubernamentales para la promoción del francés y la exclusión de cualquier otra lengua, esta situación todavía continuó hasta finales del siglo XX. Por estas circunstancias, cada vez menos personas hablaban otra lengua distinta del francés y los hablantes de estas lenguas solo las han usado entre amigos y familiares, por lo que, en ámbitos más formales, la comunicación solo se lleva a cabo en francés. Por lo tanto, las publicaciones en otras lenguas son muy escasas y la mayoría de estas fueron redactadas hace siglos, cuando estas lenguas tenían mayor aceptación.

---

<sup>88</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

<sup>89</sup> [https://slmc.uottawa.ca/?q=french\\_history](https://slmc.uottawa.ca/?q=french_history)

En la elección de las instancias de las lenguas habladas en Francia se tomaron en cuenta solo las más habladas que son: bretón, catalán, corso, francés, occitano y vasco; debido a hay un mayor número de textos en estas lenguas y la elección de otra lengua menos utilizada supondría mayores dificultades para su obtención. A parte de las lenguas ya mencionadas, se obtuvieron textos en el idioma provenzal ya que actualmente es considerado un dialecto del occitano; estas dos lenguas junto con el catalán pertenecen a las lenguas de *Oc*. Se tenía contemplado obtener textos en lenguas de *Oïl*, como el picardo, valón y arpitano, sin embargo, hay muy pocos digitalizados disponibles en internet y solo se pudo obtener en francés que pertenece a este grupo.

La obtención de las instancias de las lenguas más habladas de Francia como el caso del francés, catalán y vasco fue sencilla, puesto que existen varias páginas de internet escritas en estas lenguas. En los casos del catalán y vasco esto se debe a que las comunidades catalanas y vascas promueven el uso de sus propias lenguas para evitar que desaparezcan.

El corso, una lengua muy parecida al italiano, se habla en la isla de Córcega que se encuentra muy cercana a Italia. Para obtener instancias de esta lengua se utilizó *Wikipedia*, fragmentos de la Biblia en idioma corso<sup>90</sup>, libros digitalizados y un blog que contiene varias historias que están escritas en este idioma.

El occitano es una de las lenguas de *Oc*, pero al ser considerada una lengua *Patois* existen muy pocos textos escritos en esta lengua y aún menos disponibles de manera digital. Para obtener textos en este idioma se recurrió al uso de *Wikipedia*, libros y textos alojados en bases de datos de universidades francesas.

El bretón es una lengua celta proveniente del actual Reino Unido que ha logrado sobrevivir durante varios siglos, sin embargo, debido a las políticas francesas de desprestigio de las lenguas vernáculas y al poco interés de las personas en conservarla, esta lengua está en peligro de desaparecer. No obstante, existen documentos que están disponibles actualmente de forma digital, la mayoría de estos son textos antiguos que están alojados en servidores de universidades francesas. Para poder obtener algunos de los textos se recurrió a

---

<sup>90</sup> [http://projetbabel.org/bibbia\\_corsa/index.php?livre=genese&chap=38](http://projetbabel.org/bibbia_corsa/index.php?livre=genese&chap=38)

buscar libros que tuvieran la palabra *Brezhoneg* (que significa bretón en bretón), en páginas como *archive.org*.

El provenzal es otra de las lenguas de *Oc*, al igual que el occitano y el catalán, sin embargo, debido a ser considerada una lengua *Patois* no es reconocida por el gobierno francés y se cataloga como un dialecto del occitano. Por este motivo los textos en esta lengua fueron más difíciles de conseguir, ya que los pocos textos en provenzal eran etiquetados como occitano en las bases de datos de universidades francesas, sin tener conocimientos de estas dos lenguas, la única manera de poder diferenciar los textos era mediante los títulos que contenían la palabra *provençau* (provenzal en provenzal).

## 8.2 Familia germánica

Las instancias para la familia indoeuropea germánica se eligieron con base en las más representativas, habladas, e influyentes, y sobre todo basándose en las clasificaciones realizadas por los lingüistas. En total se escogieron a nueve lenguas: gótico, danés, islandés, noruego (*nynorsk* y *bokmål*), sueco, alemán, inglés y neerlandés.

Los lingüistas han clasificado a las lenguas de la familia indoeuropea germánica en tres grupos o subfamilias: Oriental, Nórdica y Occidental.

Para la familia oriental se eligió al único representante de esta, que es el idioma Gótico. A pesar de que esta lengua está muerta o extinta, se han preservado varios escritos, uno de los cuales es la traducción que se realizó de la Biblia en el siglo IV por el obispo godo *Wulfila* quien creó el alfabeto gótico o ulfilano basado en las runas y en los alfabetos griegos y latinos. Este alfabeto fue originalmente creado para escribir la Biblia en el idioma gótico.<sup>91</sup>

Actualmente existen traslitteraciones (transcripciones a un distinto alfabeto) de textos góticos en alfabeto latino para poder estudiar con mayor facilidad y hasta difundir el idioma gótico. En Internet existen distintas páginas en las que se pueden encontrar escritos en gótico utilizando el alfabeto latino, desde páginas de noticias y discusión, páginas con escritos originales de los godos y páginas donde se dan cursos básicos de gótico. Una de las páginas

---

<sup>91</sup> <http://www.wulfila.be/gothic/gotica/>

más útiles para obtener texto en gótico fue <http://www.wulfila.be/> donde se encuentran fragmentos de la Biblia escrita en gótico utilizando el alfabeto latino.

Para la familia germánica nórdica o escandinava se escogieron los estándares lingüísticos actuales que son el danés, sueco, noruego en sus dos variantes (*nynorsk* y *bokmål*) y el islandés, pero se excluyó al feroés (hablado en las islas Feroe) ya que es muy difícil encontrar textos en este idioma.

El danés es el más influyente de todas las lenguas nórdicas, ya que el imperio danés abarcó a los territorios que corresponden actualmente a Noruega, Dinamarca, Islas Feroe, islas *Shetland* y Groenlandia. Por esta razón el feroés y el noruego están basados en el danés y todos son altamente inteligibles entre sí. Para obtener textos en danés se utilizaron principalmente artículos alojados en *Wikipedia*.

Como se mencionó anteriormente, existen dos variantes de noruego que son el *nynorsk* y *bokmål*. Ambos idiomas están basados en el danés y surgieron a partir de la independencia de Noruega en el año 1814. Se decidió crear un idioma para reafirmar este evento, sin embargo, surgieron dos estándares de este, el *Nynorsk*<sup>92</sup> y el *Bokmål*<sup>93</sup>. Este último es actualmente el más utilizado en Noruega.

Para obtener las instancias de los dos estándares del noruego se obtuvieron textos en *Wikipedia* donde se tienen páginas distintas para cada uno haciendo la colección menos complicada, además se utilizaron páginas de noticias en noruego donde se pueden encontrar noticias en ambas variantes. Para poder diferenciarlas se utilizaron palabras clave que difieren en ambas como lo son: *eit, desse, Noreg, kva, heiter, ein, fargar, Regnbogen, gut, guten, gutar, gutane* que pertenecen al *nynorsk* y las palabras correspondientes al *bokmål*: *et, disse, Norge, hva, heter, en, farger, Regnbuen, gutt, gutten, gutter, guttene*. Esto se hizo así ya que los traductores convencionales no hacen una diferenciación entre ambos ya que son casi idénticos y reconocen a palabras como *kva* y *hva* como sinónimos, aunque en realidad pertenezcan a diferentes estándares.

---

<sup>92</sup> El *Nynorsk* (traducido como nuevo noruego), fue la obra de Ivar Aasen y es la mezcla del danés y lenguas nativas de Noruega.

<sup>93</sup> El *Bokmål* (traducido como idioma del libro), atribuido principalmente a Knud Knudsen como su autor, está basado en el danés.

El sueco es muy parecido al danés ya que comparten historia y costumbres. De hecho, Suecia y Dinamarca son países vecinos separados únicamente por el estrecho de *Kattegat*. Además, ambos idiomas provienen de la lengua hablada originalmente por los vikingos. Las instancias del danés, noruego y sueco se obtuvieron a partir de *Wikipedia*.

El islandés es la lengua nórdica menos influenciada por otras, debido principalmente a su aislamiento geográfico. Además, es la más parecida al nórdico antiguo que a su vez fue el más parecido al gótico. A pesar de su aislamiento el islandés es inteligible entre los hablantes de otras lenguas nórdicas.

Para obtener las instancias del islandés inicialmente se utilizó *Wikipedia*, pero el número de artículos y la extensión de cada uno es muy limitado. De modo que se optó por obtener textos adicionales de páginas de noticias, donde existen cientos de artículos y la gente es muy activa al dar sus opiniones sobre las noticias que les resultan de mayor interés.

Para la familia germánica occidental solamente se tomaron en cuenta el inglés, alemán y el neerlandés. Se tuvieron que descartar otros idiomas no oficiales, de los cuales hay muy pocos documentos escritos, pues solo son utilizados de forma hablada de manera informal (y por lo tanto utilizan a los idiomas oficiales, como el alemán, para los documentos escritos). Hay otros idiomas que son reconocidos oficialmente como el caso del afrikáans, sin embargo, la adquisición de textos en este idioma no era suficiente para incluirlo en el trabajo. Por último, también existe el yidis que es una lengua germánica bastante parecida al alemán, pero utiliza al alfabeto hebreo para su escritura y las transliteraciones automáticas del alfabeto hebreo al alfabeto latino no son confiables ni siguen un estándar.

El neerlandés, alemán e inglés son los idiomas más influyentes entre las lenguas germánicas. Siendo el inglés el idioma de mayor importancia en mundo actualmente, pero es el idioma “menos germánico” de todos debido al contacto que tuvo con el latín y el francés durante mucho tiempo. Debido al prestigio de esta familia en Internet existen millones de páginas escritas en estos tres idiomas por lo que la obtención de las instancias no presentó dificultades.

## 9. Resultados

Dado un conjunto de textos (de la misma familia de lenguas), se obtiene una matriz de distancias entre lenguas comparándolas por pares, con esta se obtienen las coloraciones para cada número de grupos y sus durezas correspondientes. A continuación, se calculan los valores de solidez y resiliencia a partir de las coloraciones. Los valores de resiliencia son los que nos indican cual es el mejor número de colores para clasificar.

En las siguientes secciones se muestran los resultados obtenidos para cada familia de lenguas.

### 9.1 Resultados para la familia francesa

A continuación, se presentan los resultados producidos por el modelo para las lenguas habladas en Francia:

| Número de colores | Dureza     | Solidez    | Resiliencia       |
|-------------------|------------|------------|-------------------|
| 1                 | 0.19748367 | 0.00940398 |                   |
| 2                 | 0.06748081 | 0.00771209 | <b>0.21938163</b> |
| 3                 | 0.0310716  | 0.0066582  | 0.15828505        |
| 4                 | 0.01517201 | 0.00577981 | 0.15197488        |
| 5                 | 0.00664604 | 0.00474717 | <b>0.21752804</b> |
| 6                 | 0.00103812 | 0.00177963 | <b>1.66750932</b> |

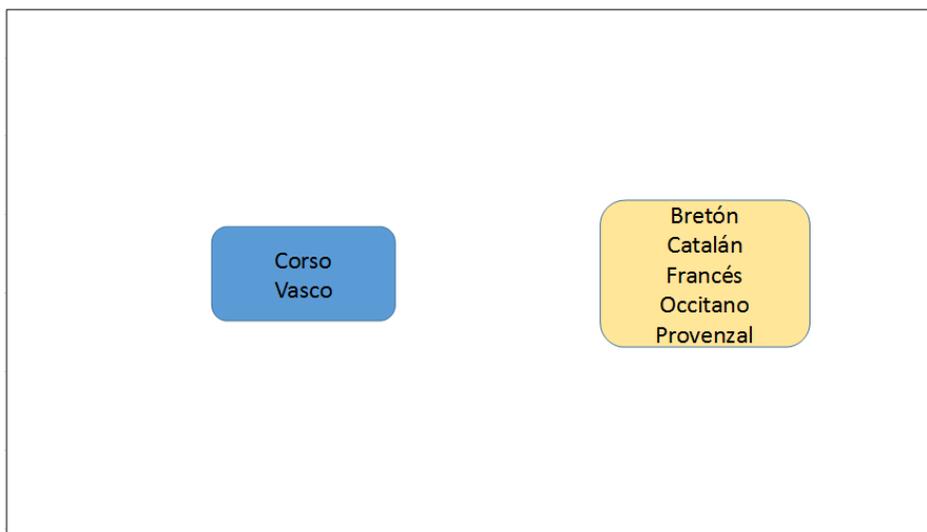
**Tabla 1.** Resultados para las lenguas de Francia

En la Tabla 1 se muestra que los valores de resiliencia mayores corresponden a las agrupaciones con dos, cinco y seis colores. La agrupación con seis colores obtuvo el mayor valor de resiliencia, por lo tanto, según el modelo, esta es la mejor clasificación para este conjunto de lenguas. En la Figura 1 se puede observar que se forma un grupo con dos lenguas que son el catalán y el occitano, mientras que las cinco lenguas restantes forman cada una un grupo distinto. Esto se debe a que el catalán y el occitano son lenguas muy parecidas y ambas han sido clasificadas como lenguas de *Oc*.



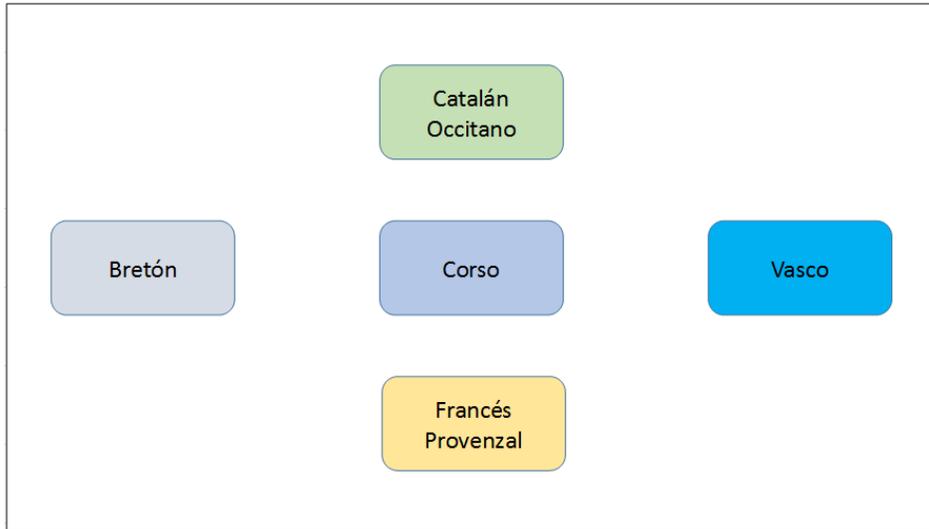
**Figura 8.** Clasificación para seis colores.

El segundo mayor valor de resiliencia corresponde a la clasificación con dos colores (Figura 2), en esta clasificación el corso y el vasco forman un grupo y los demás forman otro. Esto se debe a que la mayoría de hablantes de vasco viven en el País Vasco, mientras que los que utilizan el corso se encuentran en la isla de Córcega. Estas localidades se encuentran muy distantes entre sí, por lo tanto, los hablantes de estas lenguas prácticamente no están en contacto, por lo que sus lenguas no se han influenciado entre sí.



**Figura 9.** Clasificación para dos colores.

La tercera clasificación que obtuvo mayor valor de resiliencia es la de cinco colores (Figura 3). En esta clasificación el catalán y el occitano forman un grupo; el francés y provenzal forman otro y los tres restantes forman tres grupos distintos. El catalán, el occitano y el provenzal son consideradas como lenguas de Oc, mientras que el francés pertenece a la subfamilia de Oïl. De esta manera el provenzal y el francés pertenecen a distintos grupos. Sin embargo, existe una lengua, llamada arpitano que podría colocarse como intermedio entre el francés y el provenzal, esta se usa entre las zonas lingüísticas de Oc y Oïl, es también llamada franco-provenzal, pero no se incluyó en este trabajo debido a la escasa cantidad de texto disponible en internet. Esto muestra que el francés y el provenzal fueron influenciadas de forma recíproca a lo largo de su historia, dada la cercanía geográfica de los hablantes de cada lengua.



**Figura 10.** Clasificación para cinco colores.

Para las clasificaciones con tres y cuatro colores, se muestra un fenómeno inusual, en ambas clasificaciones se encuentran en un mismo grupo el catalán, el occitano y el francés; Los primeros dos idiomas pertenecen al grupo de Oc, mientras que el francés corresponde al de Oïl. La única explicación lógica es que los textos tomados del catalán y occitano son más recientes que las otras lenguas regionales y por tanto se muestran más influenciados por el francés a causa de las leyes que promueven el uso de este en todo el territorio.

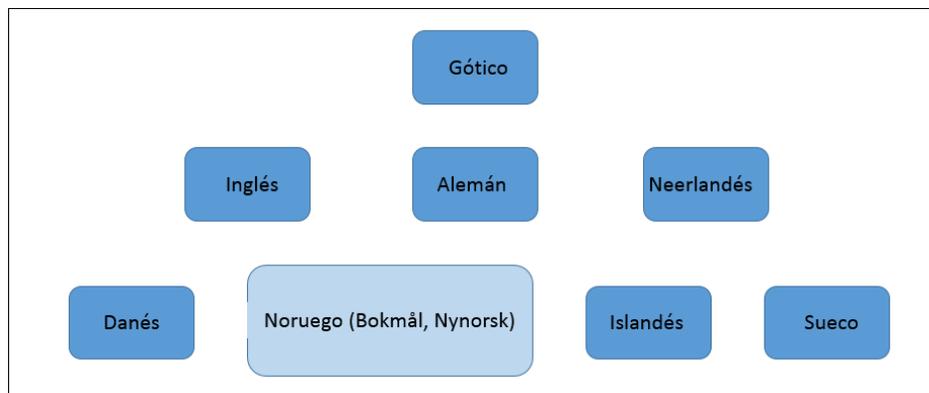
## 9.2 Resultados para la familia germánica

En la siguiente tabla se muestran los resultados obtenidos para las lenguas germánicas:

| Número de colores | Dureza  | Solidez    | Resiliencia       |
|-------------------|---------|------------|-------------------|
| 1                 | 0.53187 | 0.01477417 |                   |
| 2                 | 0.13229 | 0.00839937 | <b>0.75896232</b> |
| 3                 | 0.05161 | 0.00573444 | <b>0.46472168</b> |
| 4                 | 0.02771 | 0.00492622 | 0.16406532        |
| 5                 | 0.01279 | 0.00355278 | 0.38658327        |
| 6                 | 0.00628 | 0.00279111 | 0.27289013        |
| 7                 | 0.00249 | 0.00193667 | 0.44119334        |
| 8                 | 0.00041 | 0.00072889 | <b>1.6570122</b>  |

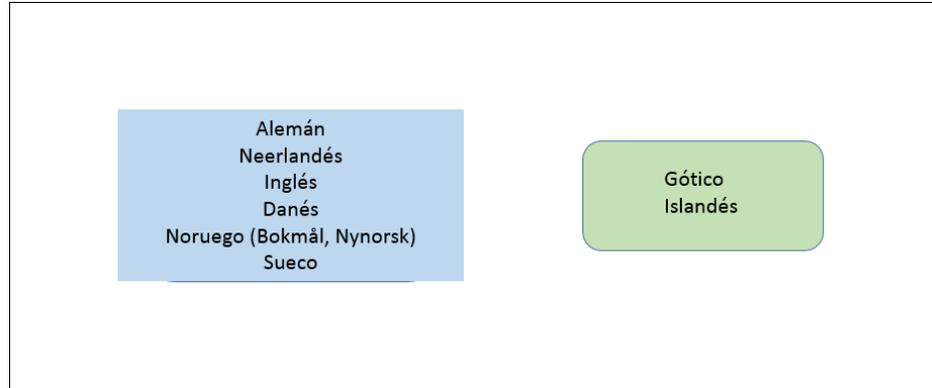
**Tabla 2.** Resultados para lenguas germánicas.

Como se observa en la Tabla 2, las mayores resiliencias corresponden a las coloraciones con dos, tres y ocho colores. La correspondiente a la de ocho colores (Figura 4) obtuvo el mayor valor, en esta asignación todas las lenguas quedan separadas en distintos grupos, excepto dos lenguas que forman uno solo, estas lenguas son los dos estándares de noruego que difieren muy poco.



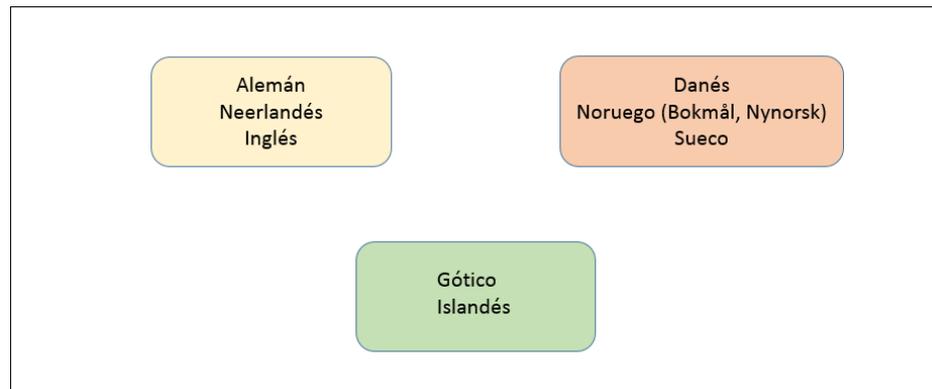
**Figura 11.** Clasificación para ocho colores.

La siguiente coloración con la mayor resiliencia es la de dos colores (Figura 5), en esta pone juntos al gótico e islandés y al resto los pone en otro grupo. Esto se puede deber a que el islandés es una lengua que se ha mantenido casi intacta durante varios siglos, es la lengua más parecida al nórdico antiguo, que a su vez era muy parecido al gótico.



**Figura 12.** Clasificación para dos colores.

La tercera coloración con mayor valor de resiliencia es la perteneciente a la de tres colores (Figura 6). Esta hace una clasificación más cercana a la propia de los lingüistas<sup>94</sup>, excepto porque al gótico e islandés el sistema los pone en un mismo grupo (por las razones ya explicadas). En el grupo que corresponde a la familia germánica occidental están el alemán, neerlandés e inglés; en el grupo que corresponde a la familia germánica oriental están el gótico e islandés (donde debería estar solo el gótico como representante de esta familia) y en el grupo correspondiente al germánico nórdico se encuentran el danés, los estándares del noruego (*nynorsk* y *bokmål*) y sueco (omitiendo al islandés).



**Figura 13.** Clasificación para tres colores.

Las demás coloraciones también son muy interesantes, ya que muestran lo que los lingüistas han dicho. Por ejemplo, en la configuración con cuatro grupos, la subfamilia

<sup>94</sup> <https://www.ethnologue.com/subgroups/germanic>

germánica occidental se presenta unida, mientras que para la de cinco colores se muestra al inglés solo en un grupo separado, ya que el inglés fue fuertemente influenciado por el francés; en la correspondiente a tres colores es parecida a la clasificación indicada por los lingüistas, pero con el islandés en el mismo grupo que el gótico, mientras que en la de cuatro colores el sueco y el islandés forman un nuevo grupo. Tal vez el sueco se separó del germánico nórdico ya que a los suecos se les dificulta mucho entender el danés y el sueco se une con el islandés ya que ambas son lenguas nórdicas. Ese cambio también indica que el danés es muy parecido a los dos estándares del noruego, esto es porque precisamente los dos estándares del noruego están basados en el estándar danés. Finalmente, en la de ocho colores los dos únicos idiomas que se encuentran en un mismo grupo son los dos estándares del noruego.

## 10. Conclusiones

En este trabajo se buscó clasificar las lenguas germánicas y francesas a partir de sus textos, utilizando un modelo de Coloración de Gráficas Suaves. Para esto se obtuvieron las frecuencias de los caracteres de cada idioma, se calcularon las diferencias de frecuencias por cada par de idiomas y estas diferencias se elevaron al cuadrado para obtener las distancias entre lenguas, es decir, las ponderaciones de cada arista de la gráfica. Una vez que se tienen las distancias, en una matriz simétrica, el sistema obtiene los valores de dureza mínimos y sus coloraciones asociadas para cada número de colores. Finalmente, con los valores de dureza se calculan los valores de solidez y resiliencia, estos últimos son los parámetros que nos indican cuales son los números de colores más apropiados para clasificar.

Este sistema arroja información muy interesante que concuerda con lo dicho por algunos lingüistas, a pesar de no coincidir con el número de grupos (subfamilias) más adecuado de acuerdo a los teóricos. Dichos lingüistas proponen una clasificación de la familia germánica que se divide en tres subfamilias (germánica oriental, nórdica y occidental).

Mientras el sistema muestra una división en tres grupos, parecida a lo propuesto en literatura especializada, basada en la coloración óptima para tres colores, con la diferencia de que el sistema pone juntos al islandés y gótico en el mismo grupo. Además, las diferencias entre las agrupaciones para cada marca de clase indican qué lenguas tienen vínculos más fuertes entre ellas. Por ejemplo, los lingüistas consideran al inglés dentro de la subfamilia germánica occidental<sup>95</sup>, pero esta lengua fue fuertemente influenciada por el francés que es una lengua “romance”, el sistema muestra que el inglés se “separa” del grupo que pertenece a la familia germánica occidental desde la coloración con cinco colores indicando que su nexo es muy débil con las otras lenguas de su subfamilia. Otro resultado interesante se puede observar a lo largo de varias coloraciones hasta la que tiene siete colores donde se pueden observar juntos al danés y los dos estándares del noruego, esto concuerda con lo que indica la historia, ya que los dos estándares de noruego están basados en el danés porque Noruega perteneció al reino danés hasta hace muy poco. Finalmente, se puede observar en la

---

<sup>95</sup> <https://www.ethnologue.com/subgroups/germanic>

coloración con ocho colores donde solo dos de las nueve lenguas, se encuentran en el mismo grupo, es decir, los dos estándares del noruego.

Mientras que las clasificaciones obtenidas para las lenguas habladas en Francia nos muestran lo distintas que son estas lenguas entre sí. Además, se consiguieron textos en pocas lenguas y la mayoría de estas son lenguas aisladas lingüísticamente hablando entre sí, incluso el vasco es una lengua que no está emparentada con ninguna otra lengua europea. Lo ideal para realizar la clasificación hubiera sido tomar la misma cantidad de textos en lenguas de *Oïl* como en lenguas de *Oc*, pero debido al imperialismo cultural francés a partir de la Revolución, muchas lenguas fueron desapareciendo paulatinamente. Algunas lenguas del norte conformarían al francés antiguo, por lo que se les considero dialectos del mismo y desaparecieron debido a la imposición del francés parisino. Por otro lado, muchas de las lenguas de *Oc* se les considera dialectos del idioma occitano a excepción del catalán, puesto que la mayoría de los hablantes de esta lengua se encuentran en territorio español, donde se les ha respetado el derecho de preservar su lengua. El bretón, a pesar de encontrarse en territorio francés, se mantuvo vivo durante muchas generaciones y sus hablantes se reusaron a dejar de hablarlo, sin embargo, la tendencia de su uso actualmente está a la baja, pues las personas más jóvenes no están interesadas en aprender una lengua que solo es hablada en Bretaña. Por último, el corso es una lengua más emparentada con el italiano, a razón de que Córcega se encuentra muy cercana a Italia, con la cual comparte historia, costumbres y cultura.

## Referencias

- Abalain, H. (2007) "Le français et les langues historiques de la France", Paris: Éditions Jean-Paul Gisserot.
- Abecassis M., Ledegen G., Zouaoui K. (2011) "La francophonie ou l'éloge de la diversité", Cambridge Scholars Publishing.
- Angoustures, A. "Historia de España en el siglo XX", Barcelona, Ariel, 1995.
- Asimov, I. (1972) "The Shapping of France", Houghton Mifflin. Harcourt Publishing Co. First Edition.
- Aydon C. (2005) "Scientific Curiosity", Constable.
- Ball, Martin J., Müller, Nicole. (2010) "The Celtic Languages", Routledge. Second Edition.
- Birmingham D. (2003) "A Concise History of Portugal" Cambridge University Press. Second Edition.
- Bopp, F. (1885). "A comparative grammar of the Sanskrit, Zend, Greek, Latin, Lithuanian, Gothic, German, and Slavonic languages." Williams and Norgate. Second Edition.
- Charles, H.G. (1907) "An introduction to vulgar latin", D.C. Heath&Co.
- Crystal, D. (2011) "Bogat: The King James Bible And The English Language", Oxford University Press.
- Crystal D. (1995) "The Cambridge Encyclopedia of the English Language" Cambridge University Press.
- D'Arco Silvio Avalle, Luciana Borghi Cedrini. (1966) "Alle origini della letteratura francese: i giuramenti di Strasburgo e la sequenza di Santa Eulalia", Giappichelli.
- Diez F. (1874) "Grammaire des Langues Romanes", Tome 1. Librairie A. Franck.
- Freeborn D (1992) "From Old English to Standard English: A Course Book in Language Variation Across Time" University of Ottawa Press.
- Gamkrelidze, T.; Ivanov, V.V. (1990). "The Early Story of Indo-European Languages". *Scientific American* **263**(3): 82-89.
- Grattan T., Hawthorne J. (1899) "Holland: the history of the Netherlands", Peter Fenelon Collier
- Elliott, J. (2002) "España en Europa: Estudios de historia comparada", Elliott Publicacions de la Universitat de València; Edición: 1.

- König E. And van der Auwera J. (1994) "Preface" In: König E. And van der Auwera J. (Eds), *Germanic Languages*. Routledge, London: ix-xiii
- Lara-Velázquez, P.; Gutiérrez-Andrade, M.A.; De-los-Cobos-Silva, S.G.; Rincón-García, E.A. (2015). "Coloración de gráficas suaves". *Revista de Matemática: Teoría y Aplicaciones*. Universidad de Costa Rica. San José **22**(2): 1-7
- Lara-Velázquez, P., de-los-Cobos-Silva, S.G., Gutiérrez-Andrade, M.A., Rincón-García, E.A., Ponsich, A., Mora-Gutiérrez R.A. (2015). "Comparative Philology among Iberian languages using Soft Graph Coloring". *Proceedings of the XVIII SIGEF CONGRESS: Scientific methods for the treatment of uncertainty in social sciences* Editorial de la Universitat Rovira i Virgili, Girona, Spain.
- Lewis M.P., *Ethnologue: Languages of the World* (2009), SIL International; 16th Edition.
- Moreno-García J (1976) "La conferencia de Berlín (1884-1885)", *Historia* 16
- Orrin W. Robinson (1992) "Old English and Its Closest Relatives: A Survey of the Earliest Germanic Languages" Routledge
- Rickard P. (2003). "A history of the french language", Routledge. Second Edition.
- Pressley, J. M. "Shakespeare's Language." Shakespeare Resource Center, en <http://www.bardweb.net/language.html>. 20/Julio/2015 11:02
- Van Gelderen, E. (2006) "A History of the English Language", John Benjamins Publishing Company.
- Walter, H. (1997). „L’aventure des mots français venus d’ailleurs“, Paris: Éditions Robert Laffont.
- Willems R (2002) "Language contact at the Romance-Germanic language border." *Multilingual Matters*
- William S-Y. Wang, James W. Minett, The invasion of language: emergence, change and death, *Trends in Ecology & Evolution*, Volume 20, Issue 5, May 2005, Pages 263-26