



**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

# Tecnológico Nacional de México

Centro Nacional de Investigación  
y Desarrollo Tecnológico

## Tesis de Maestría

Método para la adquisición y reforzamiento de  
vocabulario para el aprendizaje de una lengua  
extranjera

presentada por

**Ing. Carlos Alberto Peralta Quintero**

como requisito para la obtención del grado de  
**Maestro en Ciencias de la Computación**

Director de tesis

**Dr. Noé Alejandro Castro Sánchez**

Codirector de tesis

**Dr. Juan Gabriel González Serna**

Cuernavaca, Morelos, México. Enero de 2022.



**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico  
Departamento De Ciencias Computacionales

Cuernavaca, Mor., **13/diciembre/2021**

OFICIO No. DCC/202/2021  
Asunto: Aceptación de documento de tesis  
CENIDET-AC-004-M14-OFICIO

DR. CARLOS MANUEL ASTORGA ZARAGOZA  
SUBDIRECTOR ACADÉMICO  
PRESENTE

Por este conducto, los integrantes de Comité Tutorial del C. **CARLOS ALBERTO PERALTA QUINTERO**, con número de control M18CE077, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **"Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera"**, y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

DR. NOÉ ALEJANDRO CASTRO SÁNCHEZ  
Director de tesis

DR. JUAN GABRIEL GONZÁLEZ SERNA  
Codirector de Tesis

DRA. ANDREA MAGADÁN SALAZAR  
Revisor

DR. MÁXIMO LÓPEZ SÁNCHEZ  
Revisor

C.c.p. Depto. Servicios Escolares.  
Expediente / Estudiante  
JGGS/ibm

**cenidet**  
Centro Nacional de Investigación  
y Desarrollo Tecnológico



Interior Internado Palmira S/N, Col. Palmira,  
C.P. 62490, Cuernavaca, Morelos

Tel. (01) 777 3 62 77 70, ext. 3201,  
e-mail [dcc@cenidet.tecnm.mx](mailto:dcc@cenidet.tecnm.mx)  
[www.tecnm.mx](http://www.tecnm.mx) | [www.cenidet.tecnm.mx](http://www.cenidet.tecnm.mx)





**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLOGICO  
NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico  
Subdirección Académica

Cuernavaca, Mor., 05/enero/2022  
No. de Oficio: SAC/01/2022  
Asunto: Autorización de impresión de tesis

**CARLOS ALBERTO PERALTA QUINTERO**  
**CANDIDATO AL GRADO DE MAESTRO EN CIENCIAS**  
**DE LA COMPUTACIÓN**  
**PRESENTE**

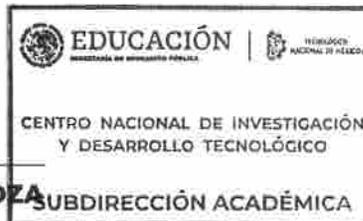
Por este conducto tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado *"Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera"*, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

**ATENTAMENTE**

*"Excelencia en Educación Tecnológica"*  
*"Educación Tecnológica al Servicio de México"*

**DR. CARLOS MANUEL ASTORGA ZARAGOZA**  
**SUBDIRECTOR ACADÉMICO**



C.c.p. Jefa(e) del Departamento de Servicios Escolares  
Expediente:  
CMAZ/CHG



Interior Internado Palmira S/N, Col. Palmira, C. P. 62490,  
Cuernavaca, Morelos Tel. (01) 777 3 62-77 73, ext. 4104,  
e-mail: acad\_cenidet@tecnm.mx  
www.tecnm.mx | www.cenidet.tecnm.mx



# Dedicatoria

Este trabajo de investigación está dedicado a mi madre, la señora Luisa Quintero Abundez, que siempre me ha apoyado en mi desarrollo académico. Partiendo con un chaleco porque su hijo saldría por primera vez en la escolta en un instituto de educación básica, hacerme ver que alguien creía en mí, hacerme saber que vale la pena aprender y hacer de uno mismo una mejor versión cada día. Y hasta mi posgrado, cuando me despertaba para llegar a tiempo a entregar ese trabajo en el que había estado trabajando arduamente y por el cual me había quedado dormido del cansancio.

No tengo dedicatoria, palabras o acciones suficientes para agradecerle todo lo que ha hecho por mí. Sin embargo, le he dicho y aquí lo escribo: *también tú cuentas conmigo por siempre.*

# Agradecimientos

Agradezco a Dios, al cielo o al universo que siempre ha puesto los medios a favor para que pueda desarrollarme personal, académica y profesionalmente.

Gracias al Gobierno de México que a través del Consejo Nacional de Ciencia y Tecnología (CONACyT) me apoyó por medio de una beca para desempeñarme como estudiante de posgrado a tiempo completo.

Un agradecimiento al TecNM/CENIDET por facilitar sus instalaciones y por las prestaciones proporcionadas durante el tiempo de mi posgrado.

Quiero expresar mi gratitud al Dr. Noé Alejandro Castro Sánchez por los conocimientos que compartió conmigo, su apoyo y tiempo, así como al Dr. Juan Gabriel González Serna por compartir su conocimiento en clases y su consejo para mejorar este trabajo de investigación. También y no menos importante, agradezco a mi comité compuesto por la Dra. Andrea Magadán Salazar y el Dr. Máximo López Sánchez por su tiempo, sus recomendaciones e ímpetu que me impidieron entrar en una zona de confort y me motivaron a entregar trabajos excelentes que rompieran con las expectativas.

Estoy agradecido también con mi padre el señor Sergio Peralta Escobar que le he aprendido buenas actitudes para lograr mi desarrollo como persona y profesionista.

Un agradecimiento a Vladimir Rodríguez Delgado que siempre estuvo apoyándome y al pendiente desde que decidí ingresar al posgrado hasta el momento cumbre de esta travesía.

Por último, quiero agradecer a mis colegas: Araceli, Rosa, Jesús, Derick, Alberto, Kenia, Karen, Juan, Julio, Magda y Neli por su amistad, apoyo y buenos momentos durante el posgrado.

# Resumen

Actualmente, los sistemas de aprendizaje de idiomas más novedosos tienen un enfoque de aprendizaje contextual o colaborativo que requiere un vocabulario amplio, con el objetivo que el estudiante realice conversaciones simuladas o con un interlocutor que domine la lengua extranjera objetivo. Sin embargo, una actividad de conversación puede resultar complicada cuando el estudiante no cuenta con el vocabulario suficiente, por tal motivo, el método de este trabajo se enfoca en actividades de adquisición y reforzamiento de vocabulario.

En esta tesis, se describe la investigación, diseño, desarrollo, implementación y evaluación de un método basado en la metodología *CALL* (*Computer Assisted Language Learning*) y centrado en el aprendizaje de vocabulario para estudiantes principiantes bajo un enfoque de aprendizaje inductivo-contextual. Este método tiene como producto final un sistema de aprendizaje de idiomas llamado *Sistema SAIC* que propone la lectura de textos cortos en lengua extranjera como fuente de vocabulario desconocido, así como una actividad didáctica de tarjetas de memoria que utiliza los fundamentos del signo lingüístico introducidos por Saussure y expuestos en Alonso (2008, pp. 91-96), con la finalidad que el estudiante aprenda y refuerce las palabras desconocidas encontradas en los textos leídos. Por otro lado, el sistema SAIC proporciona recomendaciones de textos de acuerdo con el vocabulario del estudiante mediante técnicas de filtrado basado en contenido y procesamiento de lenguaje natural (PLN), también realiza la gestión del estudio de tarjetas de memoria por medio de la técnica de repetición espaciada que tiene como fundamento la teoría de la curva del olvido de Ebbinghaus (1885). Por último, se ha evaluado el sistema SAIC con un estudio cuasi experimental longitudinal con diseño *pretest-postest* con grupo de control no equivalente, los resultados del estudio indicaron que el grupo experimental tuvo un mejor aprovechamiento en cantidad de palabras en proceso de aprendizaje en comparación con el grupo de control; sin embargo, por medio de una prueba de hipótesis de la diferencia de medias, se determinó que la diferencia observada en el aprovechamiento de los grupos estudiados no es estadísticamente significativa, este resultado se puede deber a una participación estudiantil pobre e inconsistente en las actividades realizadas durante la evaluación.

# Abstract

Currently, the most novel language learning systems carry a contextual or collaborative learning approach that requires a wide vocabulary, aiming for the student to achieve simulated conversations, or with a speaker who can be fluent in the target language. However, a conversation activity may turn out to be complicated when the student lacks of such vocabulary. Therefore, the method of this work focuses on acquisition and reinforcement vocabulary activities.

In this work is described the research, design, development, execution, and evaluation of a method based on the CALL (Computer Assisted Language Learning) methodology and focused on the vocabulary learning for beginners under an inductive-contextual learning approach. This method holds as a final output a language learning system called *SAIC System*, that suggests the reading of short texts in the foreign language as a source of unknown vocabulary, as well as a didactic activity of flashcards that deploys the principles of the Linguistic sign introduced by Saussure and exposed in Alonso (2008, pp. 91-96), in order to allow the student's learning and enhance the unknown words found in the read texts. On the other hand, the SAIC system provides texts recommendations according to the student's vocabulary via Content-based filtering techniques and Natural language processing (NLP), it also manages the study of flashcards through the Forgetting curve theory by Ebbinghaus (1885). Finally, the SAIC System has been evaluated using a longitudinal quasi-experimental study with non-equivalent control group pretest-posttest design. The study's results revealed that the experimental group obtained a better achievement in the quantity of words in learning process, opposite to the group control; however, a hypothesis test for a difference in two populations means concluded that the detected difference in the achievement of the groups is not statistically significant. This result is due to the poor and inconsistent student's participation in the performed activities throughout the evaluation.

# Contenido

<b>Capítulo 1 Introducción</b>	<b>1</b>
1.1 Descripción del problema	2
1.2 Justificación	3
1.3 Objetivos	4
<i>1.3.1 Objetivo general</i>	4
<i>1.3.2 Objetivos específicos</i>	4
1.4 Alcances y limitaciones	5
<i>1.4.1 Alcances</i>	5
<i>1.4.2 Limitaciones</i>	5
1.5 Organización del documento	6
<b>Capítulo 2 Marco conceptual</b>	<b>7</b>
2.1 Aprendizaje electrónico	7
2.2 Aprendizaje de idiomas asistido por computadora	7
2.3 La memoria	7
2.4 Sistemas de recomendación	7
2.5 Problema del inicio en frío	8
2.6 Lenguaje natural	8
2.7 Procesamiento de lenguaje natural	8
2.8 Preprocesamiento	8
2.9 Normalización	9
2.10 Tokenización	9
2.11 Lematización	9
2.12 Vectorización	9
2.13 Espacio vectorial	9
2.14 FreeLing	10
2.15 Palabras vacías	10
2.16 Repetición espaciada	10

2.17 Tarjetas de memoria	10
2.18 Vocabulario	11
2.19 Signo lingüístico	11
<b>Capítulo 3 Estado del arte</b>	<b>12</b>
3.1 Productos relacionados	12
3.2 Trabajos relacionados	15
3.2.1 <i>Aprendizaje asistido por computadora</i>	15
3.2.1.1 Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación (Consejo de Europa, 2001, pp. 9-21)	15
3.2.1.2 Del enfoque basado en tareas a la perspectiva de la cooperación (Puren, 2004)	15
3.2.1.3 El enfoque inductivo-contextual de "La Lengua Latina ilustrada por sí misma" por Hans Henning Orberg (Beccari & Penavel, 2014)	16
3.2.1.4 Estrategias pedagógicas de aprendizaje electrónico en la enseñanza médica tradicional (Wei, et al., 2018)	16
3.2.1.5 Importancia de las estrategias de autorregulación en el aprendizaje y sus derivados para la enseñanza. Análisis de un caso de Educación Superior Universitaria (Zangara & Sanz, 2015)	16
3.2.1.6 Diseño y evaluación de una aplicación tecnológica para la enseñanza del español como lengua extranjera (Ferreira, Vine, & Elejalde, 2015)	17
3.2.1.7 Estilos de aprendizaje VAK en estudiantes de Educación Física y otras pedagogías en la Universidad Internacional SEK de Chile (Flores & Maureira, 2015)	17
3.2.1.8 Modelo computacional basado en teorías de desarrollo del lenguaje para el aprendizaje y entrenamiento de idiomas: módulo de vocabulario (Moreno, Narvaez, Sastoque, & Garnica, 2016)	17
3.2.1.9 Aprendizaje incidental de vocabulario en la adquisición de un segundo idioma: una revisión de la literatura (Restrepo Ramos, 2015)	18
3.2.1.10 La importancia del vocabulario en el aprendizaje de un idioma y cómo enseñarlo (Alqahtani, 2015)	18
3.2.1.11 Un sistema de aprendizaje de vocabulario interactivo basado en una lista de frecuencia de palabras y la curva del olvido de Ebbinghaus (Zeng & Lin, 2011)	18

3.2.1.12 Diseño y desarrollo del sistema de aprendizaje de vocabulario para estudiantes de un segundo idioma usando la asociación de palabras (Chalermbuntai, Kittipol, Ranong, & Tangworakitthaworn, 2017)	19
3.2.1.13 Diseño de un sistema móvil de micro-vocabulario en inglés con base en Ebbinghaus (Zheng, 2015)	19
3.2.1.14 Aprendizaje en línea de vocabulario en inglés con diferentes sistemas para personas no angloparlantes (Al-Shumari & Bella, 2014)	19
<i>3.2.2 Repetición espaciada y la teoría de la curva del olvido</i>	<i>20</i>
3.2.2.1 Usando repetición espaciada y el paradigma de videojuegos para mejorar la divulgación científica en los alumnos de jardín de niños a 12º grado de educación elemental con lecturas cortas bajo demanda en dispositivos móviles (Yeh, Toshtzar, Guertin, & Yan, 2016)	20
3.2.2.2 Diseño basado en tarjetas combinado con repetición espaciada: una nueva interfaz para mostrar elementos de aprendizaje y mejorar la recuperación activa (Pham, Chen, Nguyen, & Hwang, 2016)	20
3.2.2.3 Implicaciones de la investigación de la memoria a corto plazo para el diseño de juegos de aprendizaje móvil basado en repetición espaciada (Schimanke, Ribbers, Mertens, & Vornberger, 2015)	20
3.2.2.4 Algoritmo de recomendación combinado basado en similitud mejorada y la curva del olvido (Li, Jin, Wu, & Chen, 2019)	21
3.2.2.5 Usando principios psicológicos de almacenamiento en memoria y preferencias para mejorar los sistemas de recomendación de música (Chmiel & Schubert, 2018)	21
3.2.2.6 Una consecuencia de la revisión en aula del contenido del curso para los estudiantes africanos de primer año de ingeniería (Swart, 2018)	21
<i>3.2.3 Recomendación de contenidos</i>	<i>22</i>
3.2.3.1 Sistemas de recomendación: principios, métodos y evaluación (Isinkaye, Folajimi, & Ojokoh, 2015)	22
3.2.3.2 EVOV: un sistema de recomendación de vídeos para apoyar el aprendizaje sostenible de vocabulario (Zhang, Jia, Zhu, & Song, 2015)	22
3.2.3.3 Sistema de recomendación de películas utilizando la técnica de filtrado colaborativo basado en elementos (Ponnamm, Deepak, Sreenivasa, Siva, & Yellamati, 2016)	23
<i>3.2.4 Comparación del estado del arte con el método propuesto</i>	<i>23</i>
<b>Capítulo 4 Método propuesto</b>	<b>28</b>
4.1 Etapa de investigación	30

4.1.1 Principios de aprendizaje y enseñanza	30
4.1.1.1 Estilos de aprendizaje del modelo VAK	30
4.1.1.2 Estrategia de enseñanza del pensamiento visual	35
4.1.2 Diseño de componentes	38
4.1.2.1 Modelo del estudiante	38
4.1.2.2 Modelo del material de estudio principal	40
4.1.2.3 Modelo del vocabulario del estudiante	41
4.1.2.4 Modelo del material de estudio secundario	42
4.1.2.5 Modelo de datos	43
4.1.2.6 Diseño del algoritmo de repetición espaciada	45
4.1.2.7 Diseño del algoritmo de recomendación de textos	49
4.1.3 Actividades de aprendizaje	54
4.1.3.1 Lectura de textos	54
4.1.3.2 Ampliación del vocabulario	55
4.1.3.3 Repaso del vocabulario	55
4.2 Etapa de desarrollo	56
4.2.1 Especificaciones técnicas del sistema SAIC	56
4.2.2 Interfaz del sistema SAIC	60
4.3 Etapa de evaluación	71
4.3.1 Plan de implementación y evaluación	71
4.3.2 Investigación experimental	72
4.3.3 Estudio cuasi experimental longitudinal con diseño pretest-postest con grupo de control no equivalente	76
<b>Capítulo 5 Evaluación</b>	<b>81</b>
5.1 Corpus de textos en inglés	82
5.2 Participantes	83
5.3 Presentación de introducción	86
5.4 Competidor más cercano del sistema SAIC	87
5.5 Actividad de exploración	89
5.6 Actividades de aprendizaje	89

5.6.1 Lectura de textos	90
5.6.2 Ampliación de vocabulario	92
5.6.3 Repaso de vocabulario	94
<b>Capítulo 6 Resultados</b>	<b>98</b>
6.1 Datos muestreados	99
6.1.1 Textos leídos de la actividad lectura de textos	99
6.1.2 Palabras agregadas al vocabulario de la actividad ampliación de vocabulario	100
6.1.3 Palabras estudiadas de la actividad repaso de vocabulario	102
6.1.4 Palabras en proceso de aprendizaje de la actividad repaso de vocabulario	104
6.2 Medida pretest-postest	106
6.3 Prueba de hipótesis de la diferencia de medias	107
6.4 Discusión	111
<b>Capítulo 7 Conclusiones</b>	<b>114</b>
7.1 Mejoras técnicas y funcionales para el sistema SAIC	115
7.2 Trabajos futuros	117
<b>Referencias</b>	<b>119</b>

# Lista de figuras

<b>Figura 1</b> Ejemplos de signo lingüístico _____	11
<b>Figura 2</b> Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera _____	29
<b>Figura 3</b> Representaciones de información del pensamiento visual _____	35
<b>Figura 4</b> Cono de la experiencia _____	36
<b>Figura 5</b> Concepto general de relación del material de estudio del método propuesto _____	37
<b>Figura 6</b> Interacción de componentes del sistema SAIC del método propuesto _____	38
<b>Figura 7</b> Diagrama simplificado del modelo del estudiante _____	39
<b>Figura 8</b> Diagrama simplificado del modelo del material de estudio principal _____	41
<b>Figura 9</b> Diagrama simplificado del modelo del vocabulario del estudiante _____	42
<b>Figura 10</b> Prototipo de tarjeta de memoria _____	43
<b>Figura 11</b> Modelo de datos del sistema SAIC _____	45
<b>Figura 12</b> Pseudocódigo que implementa el algoritmo de repetición espaciada _____	49
<b>Figura 13</b> Representación trigonométrica de la similitud coseno para vectores TF-IDF _____	51
<b>Figura 14</b> Interacción entre procedimientos de las fases del algoritmo de recomendación de textos _____	53
<b>Figura 15</b> Interacción de las actividades del estudiante con otros componentes del sistema SAIC _____	54
<b>Figura 16</b> Configuración de librerías de estilos _____	58
<b>Figura 17</b> Configuración del archivo Gemfile para la gema pagy _____	59
<b>Figura 18</b> Configuración del archivo Gemfile para la gema devise _____	59
<b>Figura 19</b> Plantilla del sistema SAIC _____	61
<b>Figura 20</b> Página de inicio de sesión del sistema SAIC _____	62
<b>Figura 21</b> Página de registro de cuenta del sistema SAIC _____	63
<b>Figura 22</b> Página del menú principal del sistema SAIC _____	64
<b>Figura 23</b> Vista de la actividad de exploración del sistema SAIC _____	64
<b>Figura 24</b> Página de la lista de lecturas del sistema SAIC _____	65
<b>Figura 25</b> Vista de lecturas recomendadas de la lista de lecturas del sistema SAIC _____	66

<b>Figura 26</b>	Página del visor de lectura del sistema SAIC _____	67
<b>Figura 27</b>	Página de nueva lectura del sistema SAIC _____	68
<b>Figura 28</b>	Vista para guardar una palabra en el vocabulario de estudio del visor de lectura del sistema SAIC _____	69
<b>Figura 29</b>	Página de la lista de tarjetas de memoria del sistema SAIC _____	70
<b>Figura 30</b>	Página de repaso de vocabulario con tarjetas de memoria del sistema SAIC _____	71
<b>Figura 31</b>	Evaluación pretest-postest _____	75
<b>Figura 32</b>	Gráfica de la ecuación de la curva del olvido con nueve repeticiones espaciadas _____	77
<b>Figura 33</b>	Gráfica de la ecuación de la curva del olvido con una repetición espaciada _____	77
<b>Figura 34</b>	Primera página del artículo “Un experimento para determinar el efecto de la curva del olvido en el aprendizaje del signo lingüístico” (Peralta-Quintero, Castro-Sánchez, González-Serna, & Nieto-Benítez, 2019) _____	80
<b>Figura 35</b>	Participación de estudiantes por género _____	83
<b>Figura 36</b>	Participación de estudiantes por edad _____	84
<b>Figura 37</b>	Participación de estudiantes por estilo de aprendizaje dominante/secundario _____	84
<b>Figura 38</b>	Participación de estudiantes por estilo de aprendizaje dominante acumulado _____	85
<b>Figura 39</b>	Participación de estudiantes por estilo de aprendizaje secundario _____	85
<b>Figura 40</b>	Participación de estudiantes por registro en el sistema SAIC durante el periodo de evaluación _____	86
<b>Figura 41</b>	Participantes en la videollamada de la presentación de introducción _____	87
<b>Figura 42</b>	Portada de la presentación de introducción durante la videollamada _____	87
<b>Figura 43</b>	Selección de palabras conocidas en la actividad de exploración _____	89
<b>Figura 44</b>	Lista de lecturas con filtros para el grupo experimental _____	91
<b>Figura 45</b>	Lista de lecturas con filtros para el grupo de control _____	91
<b>Figura 46</b>	Visor de lectura para ambos grupos de estudio _____	92
<b>Figura 47</b>	Vista para guardar una palabra en el vocabulario para el grupo experimental _____	93
<b>Figura 48</b>	Vista para guardar una palabra en el vocabulario para el grupo de control _____	93
<b>Figura 49</b>	Lista de tarjetas de memoria con filtros para el grupo experimental _____	95
<b>Figura 50</b>	Lista de tarjetas de memoria con filtro para el grupo de control _____	96
<b>Figura 51</b>	Tarjeta de memoria con aprendizaje asociativo para el grupo experimental _____	96
<b>Figura 52</b>	Tarjeta de memoria con aprendizaje tradicional para el grupo de control _____	97

<b>Figura 53</b> Participación del grupo experimental en la actividad lectura de textos _____	99
<b>Figura 54</b> Participación del grupo de control en la actividad lectura de textos _____	99
<b>Figura 55</b> Cantidad promedio de textos leídos por el grupo experimental _____	100
<b>Figura 56</b> Cantidad promedio de textos leídos por el grupo de control _____	100
<b>Figura 57</b> Participación del grupo experimental en la actividad ampliación de vocabulario	101
<b>Figura 58</b> Participación del grupo de control en la actividad ampliación de vocabulario __	101
<b>Figura 59</b> Cantidad promedio de palabras agregadas al vocabulario por el grupo experimental _____	102
<b>Figura 60</b> Cantidad promedio de palabras agregadas al vocabulario por el grupo de control _____	102
<b>Figura 61</b> Participación del grupo experimental estudiando palabras de la actividad repaso de vocabulario _____	103
<b>Figura 62</b> Participación del grupo de control estudiando palabras de la actividad repaso de vocabulario _____	103
<b>Figura 63</b> Cantidad promedio de palabras estudiadas por el grupo experimental _____	104
<b>Figura 64</b> Cantidad promedio de palabras estudiadas por el grupo de control _____	104
<b>Figura 65</b> Estudiantes del grupo experimental que lograron palabras en proceso de aprendizaje _____	105
<b>Figura 66</b> Estudiantes del grupo de control que lograron palabras en proceso de aprendizaje _____	105
<b>Figura 67</b> Cantidad promedio de palabras en proceso de aprendizaje por el grupo experimental _____	106
<b>Figura 68</b> Cantidad promedio de palabras en proceso de aprendizaje por el grupo de control _____	106
<b>Figura 69</b> Resolución de la medida pretest-postest _____	107
<b>Figura 70</b> Hipótesis nula y alternativa _____	108
<b>Figura 71</b> Tabla de valores de probabilidad acumulada para la distribución normal estándar (Vargas, 2010) _____	110
<b>Figura 72</b> Gráfica de la posición de $z_c$ y $z_t$ en la curva de distribución normal estándar ____	110
<b>Figura 73</b> Histórico de palabras en proceso de aprendizaje del grupo experimental _____	112
<b>Figura 74</b> Histórico de palabras en proceso de aprendizaje del grupo de control _____	112

# Lista de tablas

<b>Tabla 1</b> Rasgos distintivos de los sistemas de aprendizaje electrónico _____	14
<b>Tabla 2</b> Trabajos sobre el aprendizaje asistido por computadora _____	24
<b>Tabla 3</b> Trabajos sobre la repetición espaciada y la teoría de la curva del olvido _____	26
<b>Tabla 4</b> Trabajos sobre la recomendación de contenidos _____	27
<b>Tabla 5</b> Cuestionario de estilos de aprendizaje del modelo VAK (Secretaría de Educación Pública, 2004) _____	31
<b>Tabla 6</b> Tabla de resultados de estilos de aprendizaje del modelo VAK (Secretaría de Educación Pública, 2004) _____	34
<b>Tabla 7</b> Análisis del material de estudio del método propuesto _____	37
<b>Tabla 8</b> Niveles de dominio de los estudiantes de una lengua extranjera del MCER (Consejo de Europa, 2001) _____	39
<b>Tabla 9</b> Comparativa de herramientas de software para clasificar textos por su complejidad de acuerdo con el MCER _____	40
<b>Tabla 10</b> Primitivas del modelo relacional y atributos personalizados _____	43
<b>Tabla 11</b> Diez primeros repasos espaciados calculados _____	47
<b>Tabla 12</b> Diez próximos repasos espaciados calculados _____	48
<b>Tabla 13</b> Especificaciones técnicas del sistema SAIC _____	56
<b>Tabla 14</b> Gráfica que determina el indicador de aprendizaje de acuerdo con el periodo disponible para la evaluación _____	78
<b>Tabla 15</b> Fuente de los textos del corpus del sistema SAIC _____	82
<b>Tabla 16</b> Cantidad de textos del sistema SAIC por complejidad de acuerdo con el MCER _____	82
<b>Tabla 17</b> Resultados de la participación del grupo experimental y de control en las actividades de evaluación _____	98
<b>Tabla 18</b> Información sobre las palabras en proceso de aprendizaje (PPA) _____	108

# Capítulo 1 Introducción

La comunicación es un proceso en el cual se transmite información de un individuo a otro. En los seres humanos esto se logra comúnmente mediante un sistema de comunicación verbal y escrito llamado idioma. A nivel mundial las actividades sociales, comerciales, políticas y científicas tienen como idioma en común al inglés; lo que permite que exista una comunicación efectiva entre individuos de diferentes naciones (Secretaría de Educación Pública, 2017). Con base en lo anterior nace la importancia del aprendizaje de dicho idioma.

El aprendizaje del inglés como lengua extranjera va tomando más fuerza entre los países no angloparlantes, sobre todo en aquellos en vías de desarrollo. En México, los esfuerzos para integrar la enseñanza de este idioma en la educación pública comenzaron en el año 2011 con la Reforma Educativa, siendo el programa nacional de inglés para la educación básica (Secretaría de Educación Pública, 2011) y el programa proyecta 100,000 para estudiantes de educación superior (Agencia Mexicana de Cooperación Internacional para el Desarrollo, 2017) las primeras iniciativas; y consolidándose en el año 2017 con la estrategia nacional de inglés (Secretaría de Educación Pública, 2017).

Actualmente, el estilo de vida de las personas está estrechamente relacionado con el uso de distintos dispositivos y tecnologías, esto genera la oportunidad de desarrollar formas de aprendizaje asistido por computadora (*e-learning*); dentro del mismo ámbito y con la necesidad de comunicarse entre personas de diferentes naciones surge el aprendizaje de idiomas asistido por computadora (*Computer Assisted Language Learning*, en inglés; o abreviado como *CALL*). Sin embargo, la mayoría de los sistemas *CALL* actuales tienen un enfoque de aprendizaje que requiere una base de vocabulario amplia, esto significa que los estudiantes principiantes deben emplear otras formas de construir su base de vocabulario.

En este trabajo se desarrolló un método basado en la metodología *CALL* que produce un sistema de aprendizaje de idiomas con un enfoque inductivo-contextual centrado en el aprendizaje de vocabulario para estudiantes principiantes mediante la recomendación de material didáctico y su periodo de estudio.

## 1.1 Descripción del problema

Desde el año 2011 México ha invertido para integrar el aprendizaje del idioma inglés como lengua extranjera al sistema de educación pública, esto se debe a la importancia que tiene preparar a los ciudadanos mexicanos para adentrarse a una sociedad y economía globalizada. No obstante, los esfuerzos no han tenido los resultados deseados de acuerdo con reportes del INEGI (Instituto Nacional de Estadística y Geografía) del año 2012 que mencionan que sólo un 9.4% de la población en México tenía conocimientos del idioma inglés, al año siguiente Consulta Mitofsky reportó un incremento al 11.6% de la población (Instituto Mexicano para la Competitividad A. C., 2015). Además, de acuerdo con EF Education First (2020), México se encuentra en la posición de 82 de 100 países en la clasificación mundial sobre el dominio y nivel de habilidad en el idioma inglés. Con base en esto, México requiere concentrar sus esfuerzos en los estudiantes principiantes de una lengua extranjera, con el objetivo de que construyan una base de vocabulario que posteriormente les permita formar estructuras más complejas del idioma inglés.

Como soporte y alternativa a la enseñanza de una lengua extranjera, las Ciencias de la Computación proponen la metodología *CALL* (*Computer Assisted Language Learning*). Sin embargo, la mayoría de los sistemas de aprendizaje de idiomas más novedosos de esta metodología tienen un enfoque de aprendizaje contextual que propone inferir una oración faltante de una conversación simulada o un enfoque de aprendizaje colaborativo que requiere mantener una conversación con otro individuo; por lo tanto, estos sistemas quedan obsoletos de cierto modo si los estudiantes son principiantes y no cuentan con una base de vocabulario suficiente. Por otro lado, los sistemas de aprendizaje de idiomas centrados en el aprendizaje de vocabulario no cuentan con las funciones necesarias para personalizar los materiales didácticos con base en el vocabulario que el estudiante está aprendiendo, por lo que se reduce la interacción con las palabras encontradas previamente y que necesitan repasarse.

Este trabajo de investigación se centró en la dificultad que tienen los estudiantes de una lengua extranjera para aprender el signo lingüístico y la forma en que las Ciencias de la Computación pueden proveer una solución.

## 1.2 Justificación

De acuerdo con la investigación de los trabajos y productos relacionados, los sistemas *CALL* centrados en el aprendizaje de vocabulario se proponen para estudiantes principiantes de una lengua extranjera, con el objetivo que dichos estudiantes formen una base de vocabulario. Estos sistemas emplean algoritmos de recomendación para personalizar el material didáctico de acuerdo con las preferencias de los estudiantes o su interacción con el sistema, es decir, los algoritmos de recomendación seleccionan un subconjunto reducido de materiales didácticos para el estudiante a partir de un conjunto más grande. Sin embargo, estos algoritmos están basados en modelos matemáticos de acuerdo con la probabilidad de que un estudiante prefiera cierto material o en teorías psicológicas del aprendizaje que recomiendan cierto material según los patrones de aprendizaje del estudiante, dejando de lado el beneficio que supone realizar el procesamiento de lenguaje natural del material didáctico disponible en el sistema y el material que ha sido repasado por el estudiante.

En este ámbito cobran relevancia los algoritmos de recomendación híbridos que suponen la unión entre modelos o técnicas que explotan sus bondades para amortiguar sus debilidades o lograr recomendaciones con necesidades particulares. Por consiguiente, un algoritmo que unifique un modelo de filtrado basado en contenido convencional con técnicas de procesamiento de lenguaje natural puede proporcionar recomendaciones de material didáctico de acuerdo con el vocabulario de próximo repaso del estudiante de una lengua extranjera, con la finalidad que el estudiante pueda repasar frecuentemente las palabras en distintos contextos y a la vez encontrar potenciales palabras relacionadas que desconozca. Además, se puede integrar un algoritmo que gestione el periodo de repaso que el estudiante debe realizar de su vocabulario de estudio, con la finalidad de recomendar el momento de estudio y qué material repasará el estudiante.

## 1.3 Objetivos

### 1.3.1 Objetivo general

Proponer un método basado en la metodología *Computer Assisted Language Learning (CALL)* que permita al estudiante adquirir y reforzar su vocabulario para el aprendizaje de una lengua extranjera a través de un sistema de cómputo que implemente tarjetas de memoria con la técnica de repetición espaciada y la recomendación de textos basada en su vocabulario de estudio.

### 1.3.2 Objetivos específicos

- Determinar el enfoque de aprendizaje del método propuesto mediante la investigación del estado del arte y el estado de la técnica.
- Seleccionar un modelo de estilos de aprendizaje y una estrategia de enseñanza compatibles con el aprendizaje de vocabulario en un sistema *CALL*.
- Desarrollar o seleccionar una herramienta de software para perfilar el material didáctico y los estudiantes de una lengua extranjera de acuerdo con el MCER (Marco común europeo de referencia para las lenguas).
- Contar con un corpus de textos en lengua extranjera, posterior a definir las características necesarias de dichos textos.
- Permitir la gestión de tarjetas de memoria a través de la implementación de un algoritmo de repetición espaciada.
- Proporcionar textos recomendados con base en el vocabulario del estudiante mediante la implementación de un algoritmo de recomendación.
- Diseñar un plan de implementación y evaluación que, mediante un estudio experimental, permita obtener datos sobre los estudiantes y su interacción con el sistema; con el objetivo de averiguar los beneficios en el aprendizaje de los estudiantes que utilizan el sistema del método propuesto.

## **1.4 Alcances y limitaciones**

### ***1.4.1 Alcances***

- El sistema computacional que se desarrolla sirve de apoyo a estudiantes de habla hispana en el aprendizaje del idioma inglés.
- El sistema proporciona textos en inglés conforme al nivel de dominio de los estudiantes de habla hispana mediante un corpus de textos en inglés clasificados en niveles de complejidad.
- El sistema implementa un algoritmo de repetición espaciada con el que se gestionan las tarjetas de memoria y el cual permite sugerir el periodo de repaso adecuado antes de que el estudiante olvide el contenido de la tarjeta.
- El sistema recomienda textos considerando el vocabulario del estudiante.
- El sistema tiene la posibilidad de ampliar su cobertura a otros idiomas siguiendo el método propuesto.

### ***1.4.2 Limitaciones***

- En este trabajo se limita el sistema al aprendizaje del idioma inglés por estudiantes de habla hispana.
- El tamaño del corpus utilizado es pequeño para un sistema productivo que requiere cantidad y variedad de material didáctico, sin embargo, es adecuado para fines académicos de evaluación del sistema de cómputo propuesto.
- El tamaño del corpus depende de la cantidad de textos encontrados que tengan una licencia que permita su uso en el ámbito académico.
- El sistema no asegura que los estudiantes de una lengua extranjera mejoren su vocabulario, la adquisición de nuevas palabras y el reforzamiento de palabras con las que ya han interactuado depende de la constancia con la que utilicen el sistema.
- El sistema no sustituye las clases tradicionales de aprendizaje de inglés debido a que se diseñó como un complemento a estas.

## 1.5 Organización del documento

Este documento presenta una estructura de seis capítulos que integran la definición de argumentos, investigación, desarrollo, evaluación, resultados y conclusión del tema de tesis. A continuación, se da una breve explicación sobre el contenido de cada capítulo.

- Capítulo 1 Introducción. Describe el origen de la investigación, el problema que se trata de solucionar y la justificación para elaborar esta tesis, así como sus objetivos, alcances y limitaciones.
- Capítulo 2 Marco conceptual. Proporciona conceptos indispensables que el lector debe conocer para tener una mejor comprensión al momento de dar lectura al presente trabajo de tesis.
- Capítulo 3 Estado del arte. Presenta un análisis sobre los productos publicados y relacionados con el tema de tesis; además, se expone la investigación en la literatura sobre el aprendizaje asistido por computadora, repetición espaciada, recomendación de contenidos y un análisis contrastando estos trabajos relacionados con la propuesta que se presenta en esta tesis.
- Capítulo 4 Método propuesto. Expone la estructura del método que se propone en la presente tesis para darle solución al problema descrito.
- Capítulo 5 Evaluación. Describe la manera en que se realizó la evaluación del método propuesto, el sistema de cómputo utilizado, los estudiantes participantes y las actividades realizadas.
- Capítulo 6 Resultados. Expone los datos muestreados en la evaluación descrita en el capítulo anterior, así como un análisis de la información obtenida y los hallazgos encontrados.
- Capítulo 7 Conclusiones. Presenta las conclusiones a las que se ha llegado mediante los hallazgos encontrados en el capítulo anterior, además, se proporciona una lista de mejoras técnicas al sistema SAIC y las recomendaciones sobre trabajos futuros.

# Capítulo 2 Marco conceptual

En este capítulo se presentan algunos conceptos básicos que el lector debe conocer antes de dar lectura a este trabajo.

## 2.1 Aprendizaje electrónico

El aprendizaje electrónico (*e-learning*) se refiere al estudio de contenidos a través de canales electrónicos como Internet, transmisión satelital, TV interactiva, CD-ROM, etc., mediante dispositivos capaces de acceder a estos canales de comunicación (Ozkan & Koseler, 2009).

## 2.2 Aprendizaje de idiomas asistido por computadora

El aprendizaje de idiomas asistido por computadora (*Computer Assisted Language Learning* o *CALL*) se refiere al uso de canales electrónicos y dispositivos de cómputo para la enseñanza y aprendizaje de una lengua extranjera (He, 2015). Este tipo de aprendizaje involucra la investigación multidisciplinaria de áreas como la computación, psicología, educación, etc., que competen para el proceso de aprendizaje de un estudiante mediante tecnologías de la información y la comunicación.

## 2.3 La memoria

Existen numerosas investigaciones sobre la memoria que la dividen en corto y largo plazo. En general, la memoria a corto plazo se describe como un almacenamiento de capacidad y duración limitada, este tipo de memoria tiene un mecanismo activo y otro pasivo: el primero representa la información mientras se trabaja con ella mientras que el segundo representa la información una vez que se deja de utilizar. Por otro lado, la memoria a largo plazo se describe como un almacenamiento relativamente permanente y con una capacidad difícil de determinar (Schimanke, Ribbers, Mertens, & Vornberger, 2015).

## 2.4 Sistemas de recomendación

Los sistemas de recomendación tratan de resolver el problema de la sobrecarga de información en los repositorios digitales. Estos sistemas filtran fragmentos de información relevante a partir de una gran cantidad de datos, el filtrado de información se realiza

comúnmente de acuerdo con las preferencias de un usuario, sus intereses o su comportamiento observado (Isinkaye, Folajimi, & Ojokoh, 2015).

## **2.5 Problema del inicio en frío**

De acuerdo con Burke citado en Isinkaye, Folajimi y Ojokoh (2015), se refiere a una situación en la que un sistema de recomendación no cuenta con la información necesaria para realizar predicciones relevantes. El problema se produce cuando el usuario de un sistema es nuevo, no ha calificado elementos o no tiene suficiente interacción para determinar sus preferencias. Este problema representa uno de los principales factores que reduce el rendimiento en los sistemas de recomendación.

## **2.6 Lenguaje natural**

El lenguaje natural se refiere al lenguaje que utilizan los seres humanos para comunicarse, también se puede definir como un conjunto de símbolos gráficos, verbales o gesticulares que se combinan para transmitir información de un individuo a otro (Chopra, Prashar, & Sain, 2013).

## **2.7 Procesamiento de lenguaje natural**

El procesamiento de lenguaje natural (PLN) es un subcampo de la inteligencia artificial (IA) y la lingüística, este subcampo se dedica a hacer que las computadoras comprendan el lenguaje humano (Chopra, Prashar, & Sain, 2013). La comprensión del lenguaje natural les permite a las computadoras realizar tareas como extracción de información, traducción automática, recuperación de información, resumen automático, etc.

## **2.8 Preprocesamiento**

El preprocesamiento es una actividad clave en los algoritmos de minería de datos orientados al análisis de textos. Por ejemplo, en un algoritmo de categorización de texto tradicional se realiza un preprocesamiento previo a los pasos de extracción, selección y categorización. El preprocesamiento consta generalmente de tareas como tokenización, filtrado, lematización y derivación (Allahyari, Pouriye, Assefi, Safaei, Trippe, Gutierrez, & Kochut, 2017).

## **2.9 Normalización**

Es una técnica de procesamiento de lenguaje natural que consiste en la reestructuración lógica de un documento de texto (Mishra & Vishwakarma, 2015). La normalización remueve elementos no textuales o no útiles con la finalidad de crear una estructura de los datos con la capacidad de búsqueda, identificación, caracterización y derivación.

## **2.10 Tokenización**

Es una técnica de procesamiento de lenguaje natural que consiste en la división de una secuencia de caracteres en fragmentos llamados *tokens* (Allahyari, *et al.*, 2017). Los *tokens* pueden contener letras, sílabas, palabras, oraciones, etc. dependiendo del análisis que se desea realizar sobre el texto.

## **2.11 Lematización**

Es una técnica de procesamiento de lenguaje natural que realiza un análisis morfológico de las palabras que consiste en convertir y agrupar las formas flexionadas de las palabras para ser representadas como un solo elemento (Allahyari, *et al.*, 2017).

## **2.12 Vectorización**

En procesamiento de lenguaje natural, se refiere a la representación de un documento de texto mediante valores numéricos llamados *vectores* (Allahyari, *et al.*, 2017). Este vector es un arreglo sucesivo de valores numéricos que pueden representar cada palabra o *token* del texto.

## **2.13 Espacio vectorial**

De acuerdo con la descripción matemática de Allahyari, *et al.* (2017), un espacio vectorial es un modelo algebraico que representa documentos en lenguaje natural de una manera formal mediante el uso de vectores en un espacio lineal multidimensional. Se puede utilizar para realizar tareas de filtrado, recuperación, indexación o clasificación de textos.

## **2.14 FreeLing**

FreeLing es una librería que proporciona servicios de análisis de lenguaje como: análisis morfológico, etiquetado PoS, clasificación de entidades nombradas, desambiguación semántica, detección de idioma, entre otros (TALP Research Center, 2021).

## **2.15 Palabras vacías**

Las palabras vacías o *stopwords* se refieren a un conjunto de palabras que no poseen un significado o relevancia para el análisis de un texto y que se encuentran muy frecuentemente en todos los textos (artículos, pronombres, preposiciones, etc.) (Mishra & Vishwakarma, 2015). En el procesamiento de lenguaje natural, se recomienda utilizar listas de *stopwords* para filtrar los textos antes del análisis, con la finalidad de evitar interferencias en el resultado o demoras en el procesamiento. No existe una lista estándar de *stopwords*, por consiguiente, en este trabajo se utiliza la lista del NLTK de Python que es compartida por Bleier (2010).

## **2.16 Repetición espaciada**

El aprendizaje mediante repetición espaciada es una técnica que consiste en realizar repeticiones de estudio a lo largo de un periodo para asimilar cierta información. Esto significa que se necesitan calcular los intervalos de tiempo más adecuados entre sesiones de estudio antes de que la información que se está estudiando se olvide (Schimanke, Ribbers, Mertens, & Vornberger, 2015). En la literatura se han encontrado tres formas básicas de implementar un mecanismo de repetición espaciada: el sistema Leitner, una ecuación basada en la teoría de la curva del olvido de Ebbinghaus (1885) y el aprendizaje de patrones de estudio de los individuos.

## **2.17 Tarjetas de memoria**

La dinámica de tarjetas de memoria consiste en colocar palabras, números, imágenes, etc. en uno o en ambos lados de un conjunto de tarjetas que se usan para estudiar y memorizar este contenido (Pham, Chen, Nguyen, & Hwang, 2016).

## 2.18 Vocabulario

El vocabulario es un conjunto de palabras y su significado, el vocabulario es un factor importante para el aprendizaje de un idioma (Chalermbuntai, Kittipol, Ranong, & Tangworakitthaworn, 2017).

## 2.19 Signo lingüístico

El signo lingüístico se describe como una asociación entre *concepto*, expresión fonética/fonológica o gráfica; e *imagen acústica*, huella psíquica de una representación del mundo exterior que es proporcionada por los sentidos de un individuo. Estos componentes que forman el signo lingüístico también se les conoce como *significante* y *significado*, respectivamente. Este concepto fue introducido por Saussure y expuesto en Alonso (2008, pp. 91-96). En la Figura 1, se muestran algunos ejemplos de signo lingüístico.

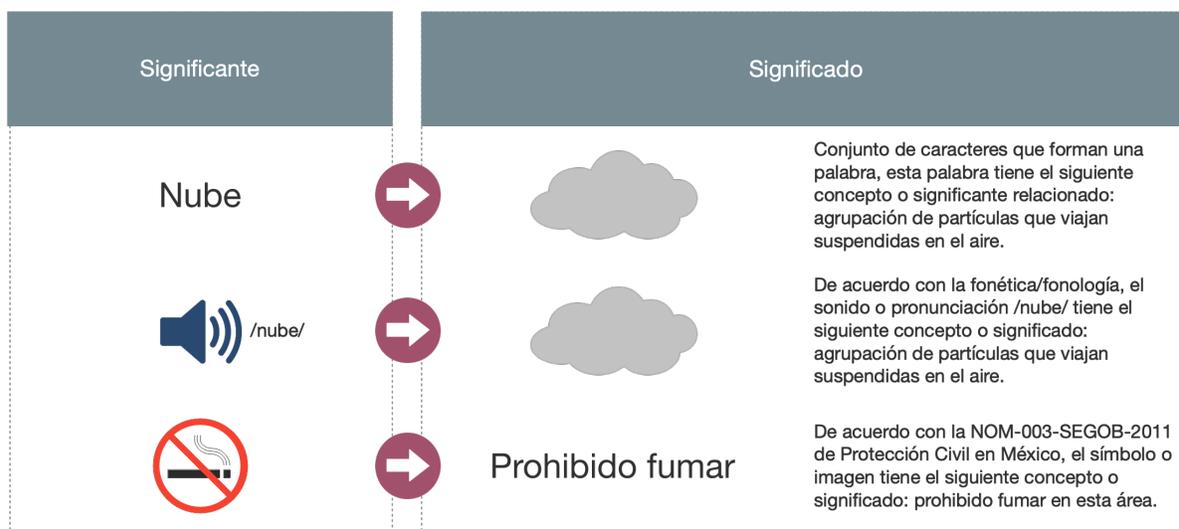


Figura 1 Ejemplos de signo lingüístico

**Nota.** El significado de la palabra *nube* es proporcionado por la Real Academia Española (2021) y el símbolo *prohibido fumar* es establecido y proporcionado por la NOM-003-SEGOB-2021 del Gobierno de México (2011).

# Capítulo 3 Estado del arte

Este capítulo se divide en dos secciones: *productos relacionados*, describe los rasgos distintivos de un conjunto de sistemas de aprendizaje de idiomas; y *trabajos relacionados*, presenta los trabajos más representativos en la revisión de la literatura. Además, las dos secciones realizan una comparativa entre los sistemas de aprendizaje de idiomas y la literatura en contraste con el método propuesto.

## 3.1 Productos relacionados

Se analizó un conjunto de sistemas de aprendizaje electrónico, con la finalidad de obtener sus rasgos distintivos. A continuación, se describen los rasgos analizados:

- Aprendizaje: indica si el sistema está orientado al aprendizaje general o de idiomas.
- Temario preestablecido: indica si el sistema cuenta con un temario de estudio.
- Paradigma de videojuegos: indica si el sistema usa elementos de videojuegos (gamificación).
- Reconocimiento de voz: indica si el sistema cuenta con tecnologías capaces de analizar o reconocer la voz.
- Recomendación de contenido: indica si el sistema es capaz de sugerir contenido personalizado al estudiante con base en sus preferencias o comportamiento observado.
- Repetición espaciada: indica si el sistema utiliza un algoritmo de repetición espaciada para programar sesiones de estudio o material didáctico.
- Fuente del material: indica la procedencia del material utilizado en el sistema de aprendizaje.
- Contenido multimedia: indica si el sistema hace uso de contenido multimedia.
- Lecturas y traducción: indica si el sistema cuenta con lectura de textos y su traducción.
- Tarjetas de memoria: indica si el sistema hace uso de tarjetas de memoria para estudiar los contenidos.

- Completa la palabra: indica si el sistema hace uso de una actividad de asociación de palabras con otras palabras u oración.
- Asociación de texto-traducción: indica si el sistema hace uso de las traducciones para asociar conceptos con el idioma que se quiere aprender.

Los sistemas de aprendizaje electrónico analizados fueron ReadLang, Anki, Quizlet, Duolingo, Busuu, Rosetta Stone, en.new with CNN, Babbel, Tandem, SuperMemo, Reverso Context, Memrise, Reading Comprehension, LinkedIn Learning, Udemy, Platzi, Código facilito, Pluralsight, Articulate, Total training y Litmos. En la Tabla 1 se muestran los 14 sistemas de aprendizaje electrónico más relevantes del conjunto analizado, incluyendo el sistema del método propuesto (sistema SAIC).

Este análisis permitió identificar los principales competidores del sistema SAIC, así como las características adicionales que presenta sobre dichos competidores. Los sistemas como Duolingo, Tandem, Rosetta Stone, Babbel, Busuu y Quizlet trabajan a nivel de oración o conversación, por lo tanto, los estudiantes principiantes que requieren formar una base de vocabulario no son su público objetivo. Por otro lado, sistemas como ReadLang, Anki, en.new with CNN, SuperMemo, Reverso Context, Memrise y Reading Comprehension trabajan a nivel de vocabulario, es decir, estos sistemas son los principales competidores del sistema propuesto. Sin embargo, el sistema propuesto destaca frente a estos competidores por la recomendación de material didáctico basado en el vocabulario del estudiante, además, las tarjetas de memoria implementan el concepto de signo lingüístico al incorporar una palabra en lengua extranjera, su pronunciación y una imagen relacionada con la palabra; con el objetivo que el estudiante construya una asociación entre dos significantes (escrito y auditivo) con un significado visual sin requerir una traducción a su lengua materna.

Tabla 1 Rasgos distintivos de los sistemas de aprendizaje electrónico

Software/Rasgo	Aprendizaje	Temario preestablecido	Paradigma de videojuegos	Reconocimiento de voz	Recomendación de contenido	Repetición espaciada	Fuente del material	Contenido multimedia	Lecturas y traducción	Tarjetas de memoria	Completa la palabra	Asociación de texto-traducción
<b>ReadLang</b> 2018/09/18	Idiomas	X	X	X	X	X	Usuarios	√ <sup>5</sup>	√	√	X	√
<b>Anki 2.0.52</b> 2018/09/18	General	X	X	X	X	√	Usuarios	X	X	√	X	X
<b>Quizlet</b> 2018/10/09	General	√	√	X	X	√	Expertos y usuarios	√ <sup>6</sup>	X	√	√	X
<b>Duolingo</b> 2018/10/09	Idiomas	√	√	X	√	√	Expertos	√ <sup>4</sup>	√ <sup>7</sup>	X	√	√
<b>Busuu</b> 2018/10/11	Idiomas	√	X	X	X	√	Expertos	√ <sup>1</sup>	X	√	√	X
<b>Rosetta Stone</b> 2018/10/12	Idiomas	√	X	√	X	√	Expertos	√ <sup>4</sup>	X	√	X	X
<b>en.new CNN</b> 2018/10/23	Idiomas	X	√	X	X	X	Repositorio	√ <sup>1</sup>	√	X	√	√
<b>Babbel</b> 2018/10/24	Idiomas	√	X	√	X	√	Expertos	√ <sup>2</sup>	√	√	√	√
<b>Tandem</b> 2018/10/25	Idiomas	X	X	X	X	X	Usuarios	√ <sup>3</sup>	√ <sup>7</sup>	X	X	X
<b>SuperMemo</b> 2018/11/15	Idiomas	√	X	X	X	√	Expertos y usuarios	√ <sup>5</sup>	√ <sup>7</sup>	X	√	X
<b>Reverso Context</b> 2018/11/15	Idiomas	X	X	X	X	X	Repositorio	√ <sup>1</sup>	√	X	√	√
<b>Memrise</b> 2018/11/15	Idiomas	√	√	X	X	√	Expertos y usuarios	X	X	√	√	√
<b>Reading Comprehension</b> 2018/11/20	Idiomas	X	X	X	X	X	Expertos	X	√ <sup>7</sup>	X	X	√
<b>Sistema SAIC</b> 2021	<b>Idiomas</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>√</b>	<b>√</b>	<b>Corpus y usuarios</b>	<b>√<sup>4</sup></b>	<b>√</b>	<b>√</b>	<b>X</b>	<b>√<sup>8</sup></b>

**Nota.**

Material de usuarios: se refiere al material que los mismos usuario suben o generan en un sistema.

Material de expertos: se refiere al material creado por expertos para la empresa a cargo del sistema.

Repositorio o corpus: se refiere al material creado por terceros y utilizado para el sistema con fines académicos.

Los sistemas LinkedIn Learning, Udemy, Platzi, CódigoFacilito, Pluralsight, Articulate, TotalTraining y Litmos fueron descartados del análisis debido a que están basados en un rasgo que no se va a analizar: los video tutoriales.

Los sistemas Kids reading comprehension 1, Reading comprehension grades 2-3 y Reading comprehension prep se omiten del análisis debido a que son idénticos al sistema Reading comprehension.

<sup>1</sup> Sólo audio

<sup>2</sup> Sólo imágenes

<sup>3</sup> Sólo videollamada

<sup>4</sup> Audio e imágenes

<sup>5</sup> Audio, vídeo y subtítulos

<sup>6</sup> Diagramas y videojuegos

<sup>7</sup> Sin traducción

<sup>8</sup> Sólo en lecturas

## **3.2 Trabajos relacionados**

El análisis de la literatura se dividió en tres apartados: 1) *Aprendizaje asistido por computadora*: consiste en la investigación sobre enfoques de aprendizaje y estrategias de enseñanza centrados en que el estudiante construya una base de vocabulario mediante sistemas de cómputo. 2) *Repetición espaciada y la teoría de la curva del olvido*: se investigaron algunos procedimientos para implementar la técnica de repetición espaciada y observar los beneficios en el aprendizaje reportados en la literatura, se encontró que la teoría de la curva del olvido de Ebbinghaus (1885) permite determinar el momento óptimo de estudio del material didáctico; es decir, planificar las repeticiones espaciadas, además de clasificar dicho material por su complejidad de aprendizaje. 3) *Recomendación de contenidos*: se investigaron los modelos que utilizan los sistemas de recomendación para realizar predicciones de contenido de acuerdo con su tipo o con su afinidad con el usuario.

### **3.2.1 Aprendizaje asistido por computadora**

#### **3.2.1.1 Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación (Consejo de Europa, 2001, pp. 9-21)**

Este documento unifica diversas directrices sobre el aprendizaje y enseñanza de una lengua dentro del contexto europeo. El documento presenta escalas descriptivas de los niveles lingüísticos de los estudiantes y los materiales didácticos. Además, propone un enfoque colaborativo que describe a los estudiantes de una lengua extranjera como agentes sociales, es decir, miembros de la sociedad que tienen actividades (fuera del aprendizaje de idiomas) que deben cumplir en determinado periodo, circunstancia y campo de acción. Por tal motivo, se orienta a construir su conocimiento sobre una lengua extranjera a través de su ambiente social y en conjunto con los individuos que lo rodean.

#### **3.2.1.2 Del enfoque basado en tareas a la perspectiva de la cooperación (Puren, 2004)**

En este trabajo, Puren (2004) realiza una comparación entre el enfoque comunicativo y el enfoque colaborativo. El enfoque comunicativo consiste en la realización de tareas para el aprendizaje de una lengua extranjera, este enfoque tiene más de dos siglos desarrollándose entre los académicos del aprendizaje de idiomas. Por otro lado, el enfoque colaborativo propuesto por

el Consejo de Europa (2001) propone un aprendizaje con base en la acción, es decir, los estudiantes adquieren habilidades, conocimientos y comprensión de una lengua extranjera en conjunto con los demás y apoyándose entre individuos que desarrollan una acción en conjunto. Sin embargo, Puren (2004) propone no considerar este último enfoque como una nueva revolución para reemplazar enfoques anteriores sino como un enriquecimiento a la colección de distintas perspectivas.

### **3.2.1.3 El enfoque inductivo-contextual de "La Lengua Latina ilustrada por sí misma" por Hans Henning Orberg (Beccari & Penavel, 2014)**

Este artículo presenta al enfoque inductivo-contextual como una propuesta para la enseñanza del latín a partir de la unión de diversos enfoques de enseñanza. De acuerdo con el enfoque inductivo-contextual, la enseñanza de una lengua extranjera debe basarse principalmente en la inferencia del estudiante por medio del contexto proporcionado.

### **3.2.1.4 Estrategias pedagógicas de aprendizaje electrónico en la enseñanza médica tradicional (Wei, et al., 2018)**

En este trabajo se presentaron distintas estrategias de aprendizaje electrónico para la enseñanza médica. Las opciones utilizadas fueron las plataformas MOOC (*Massive Open Online Courses*), aprendizaje con dispositivos móviles y Webcast (transmisión de la clase en tiempo real). Los autores concluyeron con la propuesta de un modelo de enseñanza alternativo basado en una combinación de estrategias de aprendizaje electrónico fundamentado en el modelo constructivista.

### **3.2.1.5 Importancia de las estrategias de autorregulación en el aprendizaje y sus derivados para la enseñanza. Análisis de un caso de Educación Superior Universitaria (Zangara & Sanz, 2015)**

En este artículo los autores definen la autorregulación y sus fases, denotan la importancia de mantener un medio adecuado para sostener la motivación de los estudiantes; su objetivo consistió en analizar al estudiante como un individuo enfocado a construir su propio conocimiento. En este aspecto, los autores concluyeron que existen variables que influyen en la construcción del conocimiento como el desvío de la atención en el resultado y no en el proceso

de la actividad, la poca o nula gestión del tiempo de trabajo y la necesidad de ser supervisados durante las actividades.

### **3.2.1.6 Diseño y evaluación de una aplicación tecnológica para la enseñanza del español como lengua extranjera (Ferreira, Vine, & Elejalde, 2015)**

En este trabajo Ferreira, Vine y Elejalde (2015) diseñaron e implementaron un modelo de la metodología *CALL* (*Computer Assisted Language Learning*), con la finalidad de generar un sistema de aprendizaje del idioma español. Además, los autores midieron el impacto en el aprendizaje de los estudiantes de español como lengua extranjera que utilizaron su sistema, concluyendo un incremento en el aprendizaje del español por parte del grupo experimental y un nivel de satisfacción positivo.

### **3.2.1.7 Estilos de aprendizaje VAK en estudiantes de Educación Física y otras pedagogías en la Universidad Internacional SEK de Chile (Flores & Maureira, 2015)**

En este artículo, los autores realizaron una comparación de estilos de aprendizaje con estudiantes de educación física, pedagogía, historia, lenguaje, inglés y párvulo (docencia para preescolar, en México). La comparación consistió en un estudio descriptivo transaccional de 306 estudiantes que realizaron una evaluación de estilos de aprendizaje dando como resultado que los estudiantes de primer año de la carrera sobresalen en los estilos visual y kinestésico; los estudiantes de cuarto año presentan un equilibrio entre los tres estilos de aprendizaje, excepto los estudiantes de educación física y párvulo que sobresalen en el estilo kinestésico. Los autores proponen realizar estas evaluaciones, con el objetivo que los docentes presenten sus clases con materiales didácticos adecuados para los estudiantes de las distintas carreras de acuerdo con su estilo de aprendizaje.

### **3.2.1.8 Modelo computacional basado en teorías de desarrollo del lenguaje para el aprendizaje y entrenamiento de idiomas: módulo de vocabulario (Moreno, Narvaez, Sastoque, & Garnica, 2016)**

Los autores proponen un modelo computacional basado en teorías de desarrollo del lenguaje para entrenar a estudiantes de idiomas. En este trabajo se implementó un módulo *Vocabulario* que consiste en la asociación de una palabra con una imagen del mundo real, esto con ayuda de un dispositivo móvil que lee un código QR colocado en un objeto del mundo real para que el

dispositivo muestre la palabra en lengua extranjera que corresponde al objeto. Los autores han concluido que los resultados de sus pruebas con el módulo *Vocabulario* son prometedoras, por lo tanto, continuarán implementando el diseño de su modelo computacional.

#### **3.2.1.9 Aprendizaje incidental de vocabulario en la adquisición de un segundo idioma: una revisión de la literatura (Restrepo Ramos, 2015)**

En este artículo se presentó una revisión de la literatura que analiza diversos estudios sobre el aprendizaje incidental de vocabulario en la adquisición de un segundo idioma. Esta revisión estudia la naturaleza del aprendizaje incidental de vocabulario a través de la lectura y las estrategias que lo promueven. El autor concluye que los estudiantes de un segundo idioma adquieren la mayoría de su vocabulario de manera incidental, es decir, mediante la lectura de textos con palabras relacionadas a las que conocen y anotaciones sobre la lectura.

#### **3.2.1.10 La importancia del vocabulario en el aprendizaje de un idioma y cómo enseñarlo (Alqahtani, 2015)**

Alqahtani (2015) presentó una investigación sobre la importancia del aprendizaje de vocabulario y algunas técnicas usadas por docentes para enseñar inglés como segundo idioma. De acuerdo con las técnicas analizadas, el autor describe al deficiente aprendizaje del signo lingüístico como el principal problema para construir una base de vocabulario. El autor concluye con la importancia de la adquisición de vocabulario en el aprendizaje de un segundo idioma y la necesidad de crear estrategias que permitan a los estudiantes aprender el correcto significado de las palabras.

#### **3.2.1.11 Un sistema de aprendizaje de vocabulario interactivo basado en una lista de frecuencia de palabras y la curva del olvido de Ebbinghaus (Zeng & Lin, 2011)**

En este trabajo se propuso un sistema de aprendizaje interactivo de vocabulario en inglés orientado al objetivo individual del estudiante y su capacidad de aprendizaje. Este sistema se basa en la clasificación de vocabulario para asignar un nivel de dificultad absoluto, es decir, una palabra muy común debería aprenderse primero, caso contrario con una palabra poco común. Los autores utilizan la teoría de la curva del olvido para que el sistema muestre al estudiante las palabras que necesita repasar antes de que las olvide.

### **3.2.1.12 Diseño y desarrollo del sistema de aprendizaje de vocabulario para estudiantes de un segundo idioma usando la asociación de palabras (Chalermbuntai, Kittipol, Ranong, & Tangworakitthaworn, 2017)**

En este artículo se presenta la aplicación *Woocabulary* que ayuda a los estudiantes de un segundo idioma a memorizar vocabulario mediante la comprensión del significado de las palabras para completar oraciones en inglés. Los autores evaluaron solamente la experiencia del usuario después de que los estudiantes utilizaran la aplicación para estudiar, los resultados fueron que los usuarios calificaron al sistema como *fácil de usar*.

### **3.2.1.13 Diseño de un sistema móvil de micro-vocabulario en inglés con base en Ebbinghaus (Zheng, 2015)**

Zheng (2015) propone un sistema móvil que apoya a dominar rápidamente el vocabulario en inglés. Este sistema se basa en la clasificación de vocabulario por complejidad mediante la teoría de la curva del olvido de Ebbinghaus (1885). El sistema propone asociar una palabra con su significado; además, presentarle al estudiante otras opciones como sinónimos, antónimos u homófonos de dicha palabra. En este trabajo se concluyó que, a diferencia del modelo tradicional, el micro-aprendizaje hace que el tiempo de aprendizaje sean flexibles y eficientes.

### **3.2.1.14 Aprendizaje en línea de vocabulario en inglés con diferentes sistemas para personas no angloparlantes (Al-Shumari & Bella, 2014)**

En este trabajo se presenta un estudio empírico que investiga y compara el aprendizaje de vocabulario utilizando diferentes sistemas de aprendizaje de idiomas basados en la Web. Este estudio tiene el propósito de medir los logros en el aprendizaje de vocabulario en personas no angloparlantes. Los autores mencionan tres tipos de sistemas basados en la Web: sistemas estáticos, sistemas adaptables y sistemas adaptados. Al-Shumari y Bella (2014) concluyen que los sistemas estáticos tienen una mayor apreciación de consistencia y orden ante los estudiantes, mientras que los sistemas adaptables y adaptados logran un mayor aprendizaje de vocabulario en inglés en los estudiantes.

### **3.2.2 Repetición espaciada y la teoría de la curva del olvido**

#### **3.2.2.1 Usando repetición espaciada y el paradigma de videojuegos para mejorar la divulgación científica en los alumnos de jardín de niños a 12° grado de educación elemental con lecturas cortas bajo demanda en dispositivos móviles (Yeh, Toshtzar, Guertin, & Yan, 2016)**

En este artículo se presenta una aplicación móvil que implementa la técnica de repetición espaciada y el paradigma de videojuegos, con el objetivo de enseñar temas de ciencia, tecnología, ingeniería y matemáticas a los alumnos desde jardín de niños hasta el 12° grado de educación elemental. Los autores analizaron dos aspectos de la repetición espaciada: momento durante una sesión de estudio y momento entre sesiones de estudio; también clasificaron el material didáctico como *material disponible* y *material en proceso de aprendizaje*. Los autores propusieron evaluar los resultados mediante un análisis de varianza (ANOVA), sin embargo, al momento de publicar este artículo no se presentaron datos de evaluación ni resultados.

#### **3.2.2.2 Diseño basado en tarjetas combinado con repetición espaciada: una nueva interfaz para mostrar elementos de aprendizaje y mejorar la recuperación activa (Pham, Chen, Nguyen, & Hwang, 2016)**

En este trabajo se presenta un análisis para diseñar tarjetas de memoria para el estudio de vocabulario tomando en cuenta los distintos tamaños de pantalla de los dispositivos móviles. Los autores utilizaron un sistema llamado *English Practice* que implementa la técnica de repetición espaciada por medio del sistema Leitner para probar el rediseño de su dinámica de tarjetas de memoria. Los autores concluyeron que los estudiantes tuvieron un mayor tiempo de uso en la aplicación, registraron mayor retención del conocimiento y preferían repasar el material de estudio con una mayor cantidad de respuestas incorrectas.

#### **3.2.2.3 Implicaciones de la investigación de la memoria a corto plazo para el diseño de juegos de aprendizaje móvil basado en repetición espaciada (Schimanke, Ribbers, Mertens, & Vornberger, 2015)**

En este trabajo se presenta una investigación sobre la interacción de la memoria a corto y largo plazo para determinar cómo se pueden enfrentar el efecto de olvido en el contexto de los juegos de aprendizaje basados en la repetición espaciada. Los autores mencionan que los intervalos de tiempo entre repeticiones de estudio pueden calcularse mediante algoritmos de

repetición espaciada que se basan en modelos matemáticos o psicológicos; y concluyen que, si el intervalo entre las sesiones de aprendizaje es demasiado largo, el material estudiado se puede olvidar y el proceso de aprendizaje debe repetirse desde el principio.

#### **3.2.2.4 Algoritmo de recomendación combinado basado en similitud mejorada y la curva del olvido (Li, Jin, Wu, & Chen, 2019)**

En este artículo se propone un algoritmo de recomendación que combina la similitud de Pearson mejorada para datos muy dispersos y la teoría de la curva del olvido de Ebbinghaus (1885) para rastrear el cambio de interés del usuario. Los autores utilizaron el conjunto de datos Movielens que maneja datos no comerciales para la recomendación de películas. Los resultados obtenidos por los autores mostraron que el algoritmo propuesto disminuye el error absoluto medio (Mean Absolute Error o MAE) en un 12%, con una cobertura promedio de 1.41% y un aumento de precisión promedio de 10.52%.

#### **3.2.2.5 Usando principios psicológicos de almacenamiento en memoria y preferencias para mejorar los sistemas de recomendación de música (Chmiel & Schubert, 2018)**

Chmiel y Schubert (2018) proponen un método basado en principios psicológicos para mejorar los sistemas de recomendación de música. Este método se compone de dos aspectos: una función basada en la teoría de la curva del olvido de Ebbinghaus (1885) sobre la retención de la información en la memoria y el modelo de U invertida de Berlyne que relaciona la familiaridad con el gusto de algún elemento. Los autores no presentan una evaluación o resultados; sin embargo, sugieren el uso de variables psicológicas y matemáticas para la recomendación de música.

#### **3.2.2.6 Una consecuencia de la revisión en aula del contenido del curso para los estudiantes africanos de primer año de ingeniería (Swart, 2018)**

El autor realizó una evaluación periódica del contenido de los cursos de sus estudiantes de primer año de ingeniería, con el objetivo de comprobar los beneficios de aplicar la teoría de Schön (1983) y Ebbinghaus (1885) sobre la memoria. Schön (1983) propone que el pensamiento práctico pasa por tres fases: el saber hacer, el pensamiento producido sobre lo que se hace según se actúa y el conocimiento como instrumento de evaluación, análisis, reconocimiento y reconstrucción. Mientras que Ebbinghaus (1885) expone su teoría de la curva del olvido que

propone que los humanos comienzan a perder los conocimientos adquiridos a través del tiempo. El autor realizó un estudio longitudinal a cuatro grupos de estudiantes de primer año de ingeniería eléctrica, los resultados fueron que la tasa promedio de aprobación después de tres semanas de revisiones distribuidas regularmente fueron del 76%; por lo tanto, el autor concluye en la recomendación a los académicos a programar al menos dos semanas de revisión en el aula.

### **3.2.3 Recomendación de contenidos**

#### **3.2.3.1 Sistemas de recomendación: principios, métodos y evaluación (Isinkaye, Folajimi, & Ojokoh, 2015)**

Isinkaye, Folajimi y Ojokoh (2015) presentan una investigación sobre las características y potencial de las diferentes técnicas de predicción en los sistemas de recomendación. Los autores mencionan tres fases generales del proceso de recomendación: a) fase de recolección de información: se puede utilizar una retroalimentación explícita, implícita o híbrida para obtener información relevante de los elementos que se desean recomendar; b) fase de aprendizaje: se emplea un algoritmo de aprendizaje o red neuronal para generar aprendizaje de los elementos disponibles o candidatos para recomendar; y c) fase de predicción: se realiza una selección de los elementos candidatos a recomendar mediante filtrado basado en contenido, colaborativo o híbrido. Por último, los autores explican que el desempeño de un algoritmo de recomendación se puede evaluar mediante su cobertura o precisión; y concluyen que los sistemas de recomendación amplían la oportunidad de recuperar información relevante y apoyan a solucionar el problema de la sobrecarga de información.

#### **3.2.3.2 EVOV: un sistema de recomendación de vídeos para apoyar el aprendizaje sostenible de vocabulario (Zhang, Jia, Zhu, & Song, 2015)**

En este artículo se presenta el sistema EVOV para apoyar el aprendizaje de vocabulario a través de vídeos. Este sistema consiste en utilizar vídeos con subtítulos para el estudio de vocabulario: el estudiante ve un vídeo, repasa su vocabulario y un algoritmo de recomendación genera un perfil de estudiante y genera predicciones sobre vídeos que pueden ser útiles para el aprendizaje del estudiante. Los autores concluyeron que el sistema EVOV proporciona recomendaciones de vídeos con base en el vocabulario que el estudiante repasa de manera eficiente y sostenida.

### **3.2.3.3 Sistema de recomendación de películas utilizando la técnica de filtrado colaborativo basado en elementos (Ponnam, Deepak, Sreenivasa, Siva, & Yellamati, 2016)**

En este trabajo se presenta un sistema de recomendación de películas utilizando un enfoque de filtrado colaborativo basado en elementos. Este sistema está fundamentado en dos principios: filtrado colaborativo basado en artículos y filtrado colaborativo basado en el usuario. Los autores proponen el siguiente procedimiento para realizar recomendaciones de películas: primero, los usuarios evalúan los elementos; después, el algoritmo de recomendación almacena las calificaciones, calcula la similitud entre las puntuaciones de los usuarios, selecciona los *k*-vecinos más cercanos, realiza una predicción de calificaciones desconocidas; por último, el sistema proporciona las recomendaciones al usuario. Los autores concluyen exponiendo los casos en que los enfoques basados en elementos pueden resultar afectados: escasez de datos, escasez de elementos calificados, usuario nuevo o un registro masivo de usuarios nuevos dentro de un corto periodo.

### ***3.2.4 Comparación del estado del arte con el método propuesto***

Este trabajo presenta un método basado en la metodología *CALL* con un enfoque de aprendizaje inductivo-contextual; aunque el MCER (Consejo de Europa, 2001) recomienda un enfoque colaborativo, Puren (2004) y otros autores no describen a este enfoque como único y revolucionario sino como una opción más dentro del aprendizaje de idiomas. Además, la literatura sobre las estrategias de enseñanza y la metodología *CALL* no definen un enfoque de aprendizaje, sino que dan apertura a nuevos mecanismos de enseñanza mediante sistemas de cómputo. Por otro lado, algunos autores recomiendan enfocar a los estudiantes con distintos materiales didácticos como una estrategia para mantener su atención, por lo que el método propuesto plantea clasificar a los estudiantes mediante estilos de aprendizaje para obtener retroalimentación sobre qué tipo de material didáctico les es más útil de acuerdo con su estilo.

En la literatura, se destaca la adquisición de vocabulario como la base del aprendizaje de una lengua extranjera e identifica a la dificultad de aprender el signo lingüístico como el mayor problema en el aprendizaje de una lengua, por lo tanto, se hace énfasis en la necesidad de desarrollar sistemas que ayuden a los estudiantes a asociar el significante con su significado. Con base en lo anterior, este trabajo se centra en el aprendizaje de vocabulario mediante la

asociación de significante y significado, es decir, el aprendizaje del signo lingüístico. Esta estrategia de enseñanza trata de evitar la traducción de palabras en lengua extranjera a lengua materna.

En la Tabla 2 se puede observar una comparativa de los trabajos del apartado *aprendizaje asistido por computadora* con el método propuesto por medio de los parámetros: título del trabajo, año, enfoque de aprendizaje, estrategias de enseñanza y método de evaluación empleado.

*Tabla 2 Trabajos sobre el aprendizaje asistido por computadora*

<b>Título del trabajo</b>	<b>Año</b>	<b>Enfoque de aprendizaje</b>	<b>Estrategias de enseñanza</b>	<b>Método de evaluación</b>
Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación (Consejo de Europa, 2001)	2001	Colaborativo	Actividad colaborativa Clasificación de estudiantes y material	Estandarizada
Del enfoque basado en tareas a la perspectiva de la cooperación (Puren, 2004)	2004	Colaborativo	Actividad colaborativa	No aplica
El enfoque inductivo-contextual de "La Lengua Latina ilustrada por sí misma" por Hans Henning Orberg (Beccari & Penavel, 2014)	2014	Inductivo-Contextual	Deductiva al contexto	No informado
Estrategias pedagógicas de aprendizaje electrónico en la enseñanza médica tradicional (Wei, et al., 2018)	2018	Constructivismo	E-Learning Tareas individuales	No informado
Importancia de las estrategias de autorregulación en el aprendizaje y sus derivados para la enseñanza. Análisis de un caso de Educación Superior Universitaria (Zangara & Sanz, 2015)	2015	Constructivismo	Autorregulación Tareas individuales	No informado
Diseño y evaluación de una aplicación tecnológica para la enseñanza del español como lengua extranjera (Ferreira, Vine, & Elejalde, 2015)	2015	Comunicativo	E-Learning Metodología CALL Tareas individuales	Estudio <i>cuasi</i> experimental sin grupo de control Evaluación de la experiencia del usuario
Estilos de aprendizaje VAK en estudiantes de Educación Física y otras pedagogías en la Universidad Internacional SEK de Chile (Flores & Maureira, 2015)	2015	No aplica	Clasificación de estudiantes	No informado
Modelo computacional basado en teorías de desarrollo del lenguaje para el aprendizaje y entrenamiento de idiomas: módulo de vocabulario (Moreno, Narvaez, Sastoque, & Garnica, 2016)	2016	Directo	E-Learning Realidad Aumentada	Evaluación de la experiencia de usuario
Aprendizaje incidental de vocabulario en la adquisición de un segundo	2015	Inductivo	Incidental Contextual	No informado

Título del trabajo	Año	Enfoque de aprendizaje	Estrategias de enseñanza	Método de evaluación
idioma: una revisión de la literatura (Restrepo Ramos, 2015)				
La importancia del vocabulario en el aprendizaje de un idioma y cómo enseñarlo (Alqahtani, 2015)	2015	No aplica	Clasificación de vocabulario	Estudio longitudinal
Un sistema de aprendizaje de vocabulario interactivo basado en una lista de frecuencia de palabras y la curva del olvido de Ebbinghaus (Zeng & Lin, 2011)	2011	Inductivo	E-Learning Contextual	No informado
Diseño y desarrollo del sistema de aprendizaje de vocabulario para estudiantes de un segundo idioma usando la asociación de palabras (Chalermbuntai, Kittipol, Ranong, & Tangworakitthaworn, 2017)	2017	Inductivo-Contextual	E-Learning Contextual	Evaluación de la experiencia de usuario
Diseño de un sistema móvil de micro-vocabulario en inglés con base en Ebbinghaus (Zheng, 2015)	2015	Inductivo	E-Learning Inductiva Relacional	No informado
Aprendizaje en línea de vocabulario en inglés con diferentes sistemas para personas no angloparlantes (Al-Shumari & Bella, 2014)	2014	No aplica	E-Learning	Estadística descriptiva Análisis de varianza (ANOVA) Prueba U de Mann-Whitney
<b>Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera (Peralta-Quintero, 2021)</b>	<b>2021</b>	<b>Inductivo-Contextual</b>	<b>E-Learning Metodología CALL Pensamiento visual Clasificación de estudiantes y material</b>	<b>Estudio cuasi experimental longitudinal pretest-postest con grupo de control no equivalente</b>

Por otro lado, este trabajo implementa la técnica de repetición espaciada mediante una ecuación propuesta por Hu y Ogihara (2011) e implementada también por Chmiel y Schubert (2018) fundamentada en los experimentos psicológicos sobre la retención de la memoria humana realizados por Ebbinghaus (1885), estos experimentos dieron origen a la *teoría de la curva del olvido* que postula que cualquier material que se esté estudiando se olvidará de forma decreciente a través del tiempo.

En la literatura se han encontrado trabajos que usan directamente la ecuación establecida por Ebbinghaus (1885), sin embargo, en este trabajo se utiliza la ecuación propuesta por Hu y Ogihara (2011) para obtener un acceso simplificado a todas las variables de la ecuación, con el objetivo de clasificar el material didáctico de acuerdo con su dificultad para cada estudiante. Además de dar seguimiento a los repasos que realiza el estudiante para calcular su aprendizaje por medio de una clasificación de vocabulario: palabras agregadas, palabras estudiadas, palabras en proceso de aprendizaje y palabras aprendidas.

En la Tabla 3 se observa una comparativa de los trabajos del apartado *repetición espaciada* y *la teoría de la curva del olvido* con el método propuesto mediante los parámetros: título del trabajo, año, implementación de repetición espaciada y método de evaluación.

*Tabla 3 Trabajos sobre la repetición espaciada y la teoría de la curva del olvido*

Título del trabajo	Año	Implementación de repetición espaciada	Método de evaluación
Usando repetición espaciada y el paradigma de videojuegos para mejorar la divulgación científica en los alumnos de jardín de niños a 12º grado de educación elemental con lecturas cortas bajo demanda en dispositivos móviles (Yeh, Toshtzar, Guertin, & Yan, 2016)	2016	Método propuesto por (Ausubel & Youssef, 1965)	Análisis de varianza (ANOVA)
Diseño basado en tarjetas combinado con repetición espaciada: una nueva interfaz para mostrar elementos de aprendizaje y mejorar la recuperación activa (Pham, Chen, Nguyen, & Hwang, 2016)	2016	Sistema Leitner (Leitner, 1972)	Método de riesgo atribuible Prueba <i>T Student</i>
Implicaciones de la investigación de la memoria a corto plazo para el diseño de juegos de aprendizaje móvil basado en repetición espaciada (Schimanke, Ribbers, Mertens, & Vornberger, 2015)	2015	Algoritmo SM2 (Wozniak, 2018)	No informado
Algoritmo de recomendación combinado basado en similitud mejorada y la curva del olvido (Li, Jin, Wu, & Chen, 2019)	2019	Teoría de la curva del olvido (Ebbinghaus, 1885)	Precisión y <i>recall</i>
Usando principios psicológicos de almacenamiento en memoria y preferencias para mejorar los sistemas de recomendación de música (Chmiel & Schubert, 2018)	2018	Ecuación propuesta por Hu y Ogihara (2011) basada en la teoría de la curva del olvido	No informado
Una consecuencia de la revisión en aula del contenido del curso para los estudiantes africanos de primer año de ingeniería (Swart, 2018)	2018	Teoría de la curva del olvido (Ebbinghaus, 1885)	Estudio longitudinal
<b>Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera (Peralta-Quintero, 2021)</b>	<b>2021</b>	<b>Ecuación propuesta por Hu y Ogihara (2011) basada en la teoría de la curva del olvido</b>	<b>Estudio cuasi experimental longitudinal pretest-postest con grupo de control no equivalente</b>

Por último, este trabajo personaliza el material didáctico mediante un algoritmo de recomendación de textos basado en el vocabulario del estudiante, este algoritmo selecciona un conjunto reducido de textos que contienen más palabras de reciente estudio y se lo proporciona al estudiante en una sección de recomendaciones. En contraste con los trabajos encontrados en la literatura, este trabajo agrega técnicas de procesamiento de lenguaje natural al filtrado basado en contenido convencional para realizar las predicciones de contenido para cada estudiante.

En la Tabla 4 se observa una comparativa de los trabajos del apartado *recomendación de contenidos* con el método propuesto mediante los parámetros: título del trabajo, año, contenido analizado, tipo de filtrado y método de análisis.

*Tabla 4 Trabajos sobre la recomendación de contenidos*

<b>Título del trabajo</b>	<b>Año</b>	<b>Contenido analizado</b>	<b>Tipo de filtrado</b>	<b>Métodos de análisis</b>
Sistemas de recomendación: principios, métodos y evaluación (Isinkaye, Folajimi, & Ojokoh, 2015)	2015	No aplica	No aplica	No aplica
EVOV: un sistema de recomendación de vídeos para apoyar el aprendizaje sostenible de vocabulario (Zhang, Jia, Zhu, & Song, 2015)	2015	Vocabulario de subtítulos de vídeos Vocabulario desconocido del estudiante	Filtrado basado en contenido	Función propuesta por los autores
Sistema de recomendación de películas utilizando la técnica de filtrado colaborativo basado en elementos (Ponnam, Deepak, Sreenivasa, Siva, & Yellamati, 2016)	2016	Artículos Usuarios	Filtrado colaborativo	K-vecinos más cercanos Similitud coseno
<b>Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera (Peralta-Quintero, 2021)</b>	<b>2021</b>	<b>Vocabulario de textos Vocabulario de próximo repaso del estudiante</b>	<b>Filtrado basado en contenido Actividad de exploración (amortigua el inicio en frío)</b>	<b>Espacio vectorial (TF-IDF) Similitud coseno</b>

## Capítulo 4 Método propuesto

El método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera se basa en la metodología *CALL* (*Computer Assisted Language Learning*), esta metodología es definida por Hubbard (1996) como un conjunto de principios bajo el contexto del uso de las computadoras y medios tecnológicos como herramienta de aprendizaje de una lengua. De acuerdo con Hubbard (1996), el modelo *CALL* se compone de tres módulos: 1) Desarrollo: se define la estrategia de enseñanza, el diseño y los procedimientos de aprendizaje del sistema de cómputo. 2) Implementación: se proponen una serie de actividades a ejecutarse antes, durante y después del uso del sistema de cómputo. 3) Evaluación: se trata de estimar el beneficio que aporta el uso del sistema de cómputo en el aprendizaje de los estudiantes de una lengua extranjera o la mejora en la experiencia del usuario.

El método propuesto integra la especificación del desarrollo de software y los módulos tradicionales de la metodología *CALL* descritos por Hubbard (1996). De acuerdo con lo anterior el módulo de desarrollo se integra en la *etapa de investigación*, la especificación del desarrollo de software se define en la *etapa de desarrollo*, por último, los módulos de implementación y evaluación se integran en la *etapa de evaluación*. De esta manera el método propuesto comienza con la investigación y análisis preliminar del estado del arte, con la finalidad de utilizar y adaptar los fundamentos del aprendizaje de idiomas asistido por computadora existentes en la literatura de acuerdo con el público objetivo y posteriormente con el enfoque de aprendizaje seleccionado. Después de la investigación preliminar, el método propone las tres etapas antes mencionadas: *investigación*, se definen los principios de aprendizaje y enseñanza, el diseño de componentes y las actividades de aprendizaje de acuerdo con la investigación preliminar realizada; *desarrollo*, se describen las tecnologías utilizadas en la construcción de un sistema Web que implementa los principios, diseños y actividades de la etapa anterior; y *evaluación*, se describen las actividades y estudio de evaluación del sistema de cómputo para obtener una estimación sobre el beneficio del sistema de cómputo en el aprendizaje de los estudiantes de una lengua extranjera, con la finalidad de mejorar el sistema o reorientar el enfoque de aprendizaje del método. La Figura 2 muestra un diagrama general del método descrito en este capítulo.

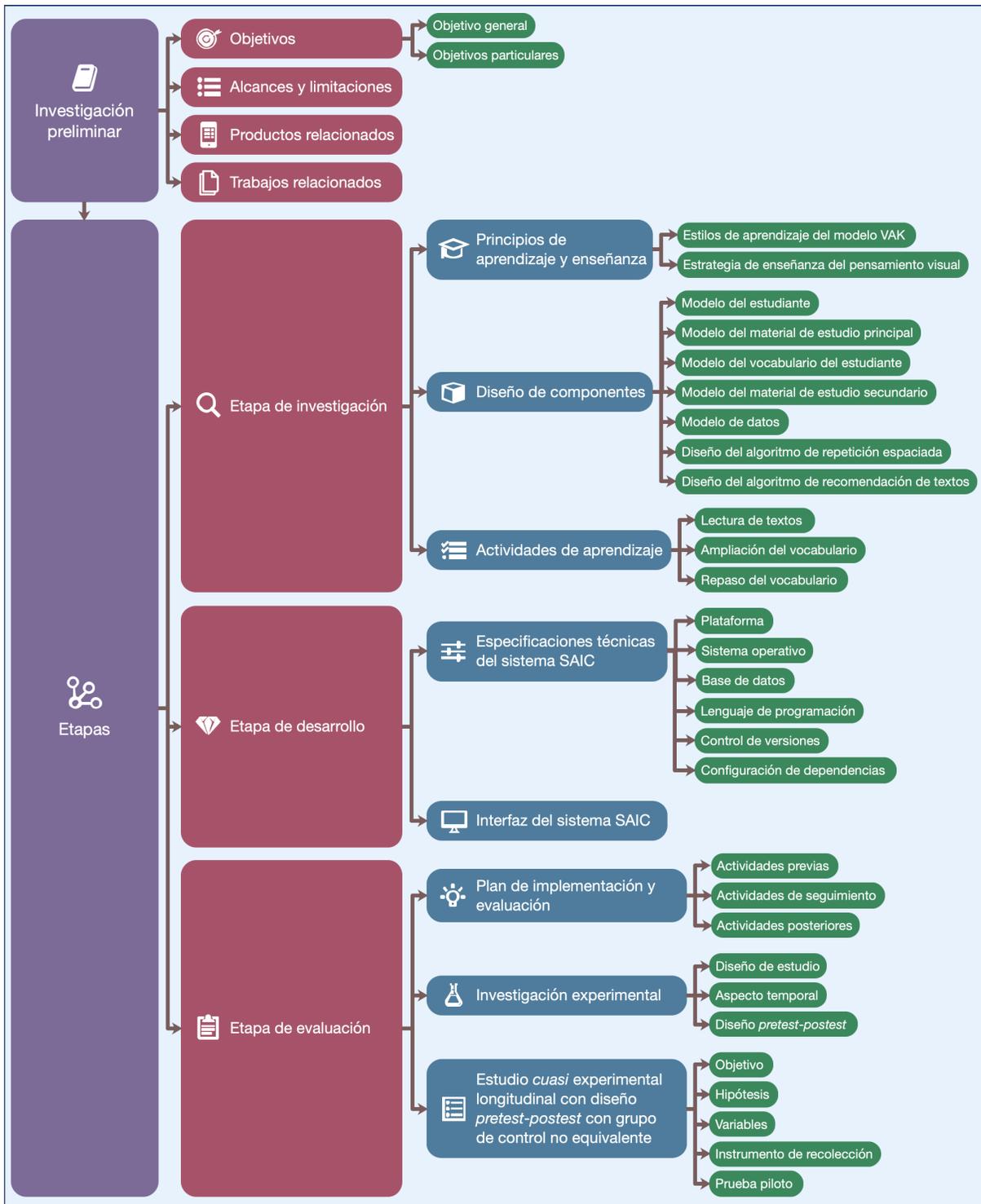


Figura 2 Método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera

## 4.1 Etapa de investigación

### 4.1.1 Principios de aprendizaje y enseñanza

De acuerdo con la literatura, los estudiantes principiantes de una lengua extranjera requieren construir una base de vocabulario y se identifica la dificultad de aprender el signo lingüístico como el mayor obstáculo para lograrla. Con base en lo anterior, este método utiliza un *enfoque inductivo-contextual* que, según Beccari y Panavel (2014), propone actividades con elementos interrelacionados, por medio de los cuales el estudiante puede deducir una palabra gracias al contexto, es decir, se plantea que el estudiante infiera un significante (palabra en lengua extranjera) mediante un significante auditivo (pronunciación de la palabra en lengua extranjera) y su significado visual (una imagen que para el estudiante mejor representa a la palabra en lengua extranjera). Esta estrategia de enseñanza trata de imitar el aprendizaje de las palabras como un niño pequeño aprendería su lengua materna, lo que trata de evitar la traducción de palabras.

La estrategia de enseñanza empleada se sustenta en el *pensamiento visual* que, en conjunto con los *estilos de aprendizaje del modelo VAK*, fomentan distintos mecanismos de aprendizaje para los estudiantes; es decir, establece las áreas de oportunidad para la implementación de módulos de software que se adapten al estudiante de acuerdo con su estilo de aprendizaje y experiencia con el sistema.

#### 4.1.1.1 Estilos de aprendizaje del modelo VAK

Un modelo de estilos de aprendizaje es un indicador sobre la manera en que los estudiantes perciben, utilizan, interpretan, construyen o resuelven la información durante su aprendizaje; este indicador puede estar basado en rasgos cognitivos, afectivos y fisiológicos (Secretaría de Educación Pública, 2004).

El método propuesto utiliza los *estilos de aprendizaje del modelo VAK* (también conocido como el modelo de programación neurolingüística de Bandler y Grinder) que posee fundamentos neurológicos y lingüísticos compatibles con las teorías sobre la memoria y la adquisición de vocabulario que se han adoptado. Este modelo considera que el ser humano puede tener hasta tres sistemas o estilos de representación mental de la información: *estilo visual*, indica que el estudiante aprende mejor mediante la lectura o cuando visualiza la

información de alguna manera (mental o gráficamente), este estilo permite extraer mucha información con gran rapidez; *estilo auditivo*, indica que el estudiante aprende de manera secuencial y ordenada mediante la expresión oral del lenguaje (sea de un expositor o de sí mismo), este estilo es indispensable para aprender algunas habilidades (como la música) aunque no es tan rápido para extraer la información o práctico para relacionarla; y *estilo kinestésico*, indica que el estudiante aprende mejor cuando relaciona la información con estímulos sensoriales o movimientos corporales, este estilo implica un aprendizaje más profundo aunque más lento para extraer la información (Secretaría de Educación Pública, 2004).

Durante la revisión de la literatura, se encontró que los estudiantes pueden tener un estilo de aprendizaje dominante con otro secundario e incluso hasta los tres estilos de aprendizaje dominantes por igual; estos casos reciben el nombre de *estilo flexible*. Además, se encontró que los estudiantes tienen una proporción general de estilos de aprendizaje dominante de 40% para el estilo visual, 30% para el estilo auditivo y 30% para el estilo kinestésico (Secretaría de Educación Pública, 2004). El perfilado de un estudiante para determinar su estilo de aprendizaje se realiza mediante un cuestionario (ver Tabla 5) y una tabla de resultados (ver Tabla 6) predefinidos por los estilos de aprendizaje del modelo VAK y proporcionados por Secretaría de Educación Pública (2004).

*Tabla 5 Cuestionario de estilos de aprendizaje del modelo VAK (Secretaría de Educación Pública, 2004)*

Pregunta	Opciones de respuesta
1. ¿Cuál de las siguientes actividades disfrutas más?	a) Escuchar música b) Ver películas c) Bailar en una fiesta
2. ¿Qué programa de televisión o de servicio de <i>streaming</i> prefieres?	a) Reportajes de descubrimientos y lugares b) Cómic y de entretenimiento c) Noticias del mundo
3. Cuando conversas con otra persona, tú:	a) Escuchas a la persona atentamente b) Observas a la persona c) Sueles tocarla (tomar su mano, brazo, hombro, etc.)
4. Si pudieras adquirir uno de los siguientes artículos, ¿Cuál elegirías?	a) Un jacuzzi b) Un reproductor de música portátil c) Un televisor
5. ¿Qué prefieres hacer un sábado por la tarde?	a) Quedarte en casa b) Ir a un concierto c) Ir al cine
6. ¿Qué tipo de exámenes se te facilitan más?	a) Examen oral b) Examen escrito c) Examen de opción múltiple

Pregunta	Opciones de respuesta
7. ¿Cómo te orientas más fácilmente?	a) Mediante un mapa b) Pidiendo indicaciones a alguna persona c) Por medio de la intuición
8. ¿En qué prefieres ocupar tu tiempo en un lugar de descanso?	a) Pensar o meditar b) Caminar por los alrededores c) Descansar
9. ¿Qué te halaga más?	a) Que te digan que tienes buen aspecto b) Que te digan que tienes un trato muy agradable c) Que te digan que tienes una conversación interesante
10. ¿Cuál de estos ambientes te atrae más?	a) Uno en el que se sienta un clima agradable b) Uno en el que se escuchen las olas del mar c) Uno con una hermosa vista al océano
11. ¿De qué manera se te facilita aprender algo?	a) Repitiendo en voz alta b) Escribiéndolo varias veces c) Relacionándolo con algo divertido
12. ¿A qué evento preferirías asistir?	a) A una reunión social b) A una exposición de arte c) A una conferencia
13. ¿De qué manera te formas una opinión de otras personas?	a) Por la sinceridad en su voz b) Por la forma de estrecharte la mano c) Por su aspecto
14. ¿Cómo te consideras?	a) Atlético b) Intelectual c) Sociable
15. ¿Qué tipo de películas te gusta más?	a) Clásicas b) De acción c) De amor
16. ¿Cómo prefieres mantenerte en contacto con otra persona?	a) Por mensajería o correo electrónico b) Tomando un café juntos c) Por teléfono
17. ¿Cuál de las siguientes frases se identifica más contigo?	a) Me gusta que mi coche se sienta bien al conducirlo b) Percibo hasta el más ligero ruido que hace mi coche c) Es importante que mi coche esté limpio por fuera y por dentro
18. ¿Cómo prefieres pasar el tiempo con tu pareja?	a) Conversando b) Acariciándose (Abrazándose, tocando su mano, etc.) c) Mirando algo juntos
19. Si no encuentras las llaves en tu bolsa o mochila, tú:	a) Las buscas mirando dentro de la bolsa o mochila b) Sacudes la bolsa o mochila para oír el ruido de las llaves c) Buscas al tacto dentro de la bolsa o mochila
20. Cuando tratas de recordar algo, ¿Cómo lo haces?	a) A través de imágenes (recordatorios de papel, lista o foto) b) A través de emociones (hilo en el dedo o algún objeto sobre tu piel) c) A través de sonidos (música, alarma o temporizador)
21. Si tuvieras dinero, ¿Qué harías?	a) Comprar una casa b) Viajar y conocer el mundo c) Adquirir un estudio de grabación
22. ¿Con qué frase te identificas más?	a) Reconozco a las personas por su voz b) No recuerdo el aspecto de la gente c) Recuerdo el aspecto de alguien, pero no su nombre

Pregunta	Opciones de respuesta
23. Si tuvieras que quedarte en una isla desierta, ¿Qué preferirías llevar contigo?	a) Algunos buenos libros b) Un radio portátil de alta frecuencia c) Golosinas y comida enlatada
24. ¿Cuál de los siguientes entretenimientos prefieres?	a) Tocar un instrumento musical b) Sacar fotografías c) Actividades manuales
25. ¿Cómo es tu forma de vestir?	a) Impecable b) Informal c) Muy informal
26. ¿Qué es lo que más te gusta de una fogata nocturna?	a) El calor del fuego y los bombones asados b) El sonido del fuego quemando la leña c) Mirar el fuego y las estrellas
27. ¿Cómo se te facilita entender algo?	a) Cuando te lo explican verbalmente b) Cuando utilizan medios visuales c) Cuando se realiza a través de alguna actividad
28. ¿Por qué te distingues?	a) Por tener una gran intuición b) Por ser un buen conversador c) Por ser un buen observador
29. ¿Qué es lo que más disfrutas de un amanecer?	a) La emoción de vivir un nuevo día b) Las tonalidades del cielo c) El canto de las aves
30. Si pudieras elegir una profesión (u otra diferente), ¿Cuál sería?	a) Un médico b) Un músico c) Un pintor de arte
31. Cuando eliges tu ropa, ¿Qué es lo más importante para ti?	a) Que sea adecuada b) Que luzca bien c) Que sea cómoda
32. ¿Qué es lo que más disfrutas de una habitación?	a) Que sea silenciosa b) Que sea confortable c) Que esté limpia y ordenada
33. ¿Qué es más sensual para ti?	a) Una iluminación tenue b) El perfume c) Cierta tipo de música
34. ¿A qué tipo de espectáculo preferirías asistir?	a) A un concierto de música b) A un espectáculo de magia c) A una muestra gastronómica
35. ¿Qué te atrae más de una persona?	a) Su trato y forma de ser b) Su aspecto físico c) Su conversación
36. Cuando vas de compras, ¿En dónde pasas mucho tiempo?	a) En una librería b) En una perfumería c) En una tienda de discos
37. ¿Cuál es tu idea de una noche romántica?	a) A la luz de las velas b) Con música romántica c) Bailando tranquilamente
38. ¿Qué es lo que más disfrutas de un viaje?	a) Conocer personas y hacer nuevos amigos b) Conocer lugares nuevos c) Aprender sobre otras costumbres

Pregunta	Opciones de respuesta
39. Cuando estás en la ciudad, ¿Qué es lo que más hechas de menos del campo?	a) El aire limpio y refrescante b) Los paisajes c) La tranquilidad
40. Si te ofrecieran uno de los siguientes empleos, ¿Cuál elegirías?	a) Director de una estación de radio b) Director de un club deportivo c) Director de un periódico o revista

**Nota.** Este cuestionario tiene el objetivo de identificar el estilo de aprendizaje de un estudiante de acuerdo con el *modelo Visual-Auditivo-Kinestésico (VAK)*, también conocido como *modelo de programación neurolingüística de Bandler y Grinder*. El estudiante debe seleccionar la opción que más lo identifique para cada una de las preguntas.

**Tabla 6** Tabla de resultados de estilos de aprendizaje del modelo VAK (Secretaría de Educación Pública, 2004)

Número de Pregunta	Visual	Auditivo	Kinestésico
1	B	A	C
2	A	C	B
3	B	A	C
4	C	B	A
5	C	B	A
6	B	A	C
7	A	B	C
8	B	A	C
9	A	C	B
10	C	B	A
11	B	A	C
12	B	C	A
13	C	A	B
14	A	B	C
15	B	A	C
16	A	C	B
17	C	B	A
18	C	A	B
19	A	B	C
20	A	C	B
21	B	C	A
22	C	A	B
23	A	B	C
24	B	A	C
25	A	B	C
26	C	B	A
27	B	A	C
28	C	B	A
29	B	C	A
30	C	B	A

Número de Pregunta	Visual	Auditivo	Kinestésico
31	B	A	C
32	C	A	B
33	A	C	B
34	B	A	C
35	B	C	A
36	A	C	B
37	A	B	C
38	B	C	A
39	B	C	A
40	C	A	B
<b>Total:</b>			

**Nota.** En la tabla de resultados se deben marcar las respuestas seleccionadas por el estudiante para cada una de las preguntas del cuestionario para identificar los estilos de aprendizaje del modelo VAK, después se realiza un conteo de las marcas obtenidas por columna y se identifica qué estilo de aprendizaje tuvo mayor puntuación. Si sólo un estilo de aprendizaje tuvo mayor puntuación, el estilo resultante es dominante. Si más de un estilo tuvo una puntuación empatada, el estilo resultante es flexible.

#### 4.1.1.2 Estrategia de enseñanza del pensamiento visual

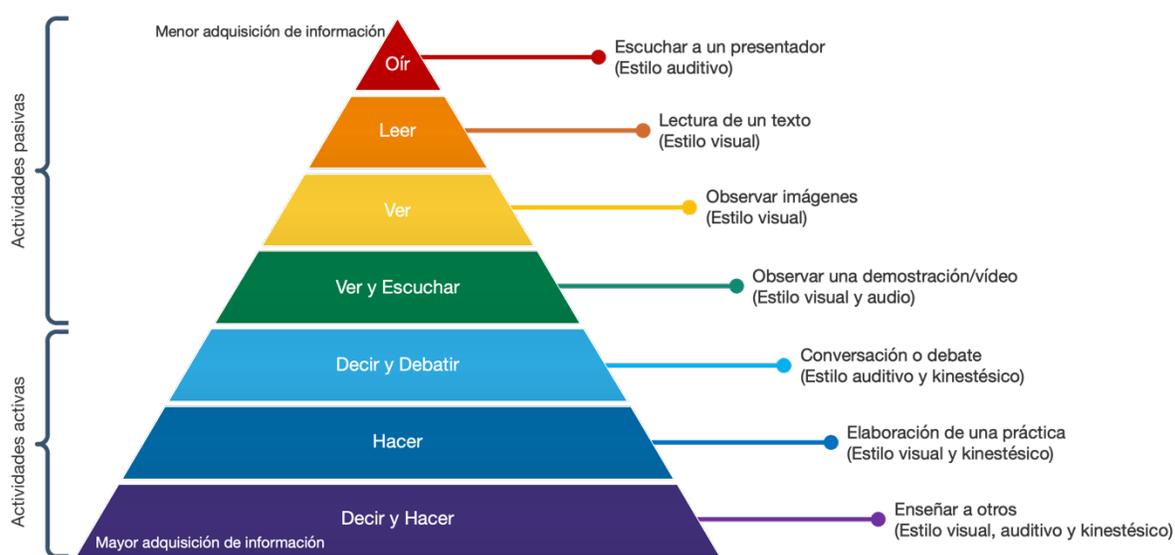
Una estrategia de enseñanza se refiere al paradigma en el que se basa un conjunto de actividades pensadas para trabajar con determinados estilos de aprendizaje. Dentro de las estrategias de enseñanza, el *pensamiento visual* consiste en crear imágenes mentales de la información mediante representaciones simbólicas, icónicas o concretas, también apoyadas de mecanismos auditivos. En la Figura 3 se describen las representaciones de información del pensamiento visual.

Representaciones		
Simbólica	1, 2, 3, A, B, C, 水	Caracteres alfanuméricos e ideogramas
Icónica		Imagen abstracta
Concreta		Imagen u objeto real

*Figura 3 Representaciones de información del pensamiento visual*

La presentación de material visual, auditivo o audiovisual siguiendo la estrategia de pensamiento visual que supone el fortalecimiento de los estilos de aprendizaje visual y auditivo cuando el estudiante presenta deficiencias en dichos estilos. Además, este material de estudio se encuentra dentro de las actividades pasivas o receptoras del aprendizaje de acuerdo con la guía psicológica del *cono de la experiencia* de Dale (1932), esta guía indica el nivel de adquisición de información a través del contenido audiovisual y de experiencia.

En la Figura 4 se puede observar el cono de la experiencia (Dale, 1932), este cono indica en su parte inferior las actividades activas o de interacción que contemplan una mayor adquisición de información mientras que en la parte superior las actividades pasivas o receptoras indican una menor adquisición de información.



*Figura 4 Cono de la experiencia*

**Nota.** Esta figura es una representación de lo expuesto por Dale (1932) sobre los métodos para analizar contenido audiovisual y de experiencia para la audiencia, además se integran los estilos de aprendizaje del modelo VAK a las actividades.

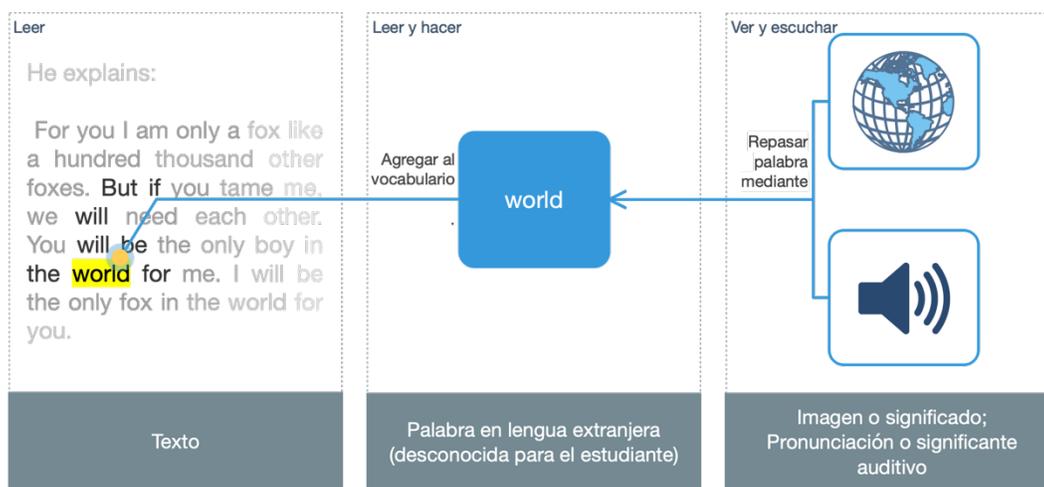
El método propuesto utiliza material de estudio como textos, significantes y su pronunciación en lengua extranjera e imágenes representativas dentro de la estrategia de enseñanza del pensamiento visual, con el objetivo que el estudiante pueda relacionar el significante con su significado, es decir, aprender los signos lingüísticos de una lengua extranjera; además se toma en cuenta el estilo de aprendizaje del modelo VAK y el nivel de experiencia del estudiante de acuerdo con el cono de Dale (1932) al que pertenece cada material.

En la Tabla 7 se muestra el material de estudio del método propuesto, su descripción, el tipo de representación del pensamiento visual, el estilo de aprendizaje al que va dirigido, la actividad y su tipo de acuerdo con la experiencia del estudiante. Además, se puede observar que el tipo de actividad pasiva es predominante, esto expone un área de mejora futura al implementar nuevas actividades activas para los estudiantes.

*Tabla 7 Análisis del material de estudio del método propuesto*

Material de estudio	Descripción	Tipo de representación	Estilo de aprendizaje	Actividad	Tipo de actividad
Texto	Texto en lengua extranjera	Simbólica	Visual	Leer	Pasiva o receptiva
Palabra en lengua extranjera	Palabra o significante en lengua extranjera que debe aprender el estudiante para construir su base de vocabulario	Simbólica	Visual y kinestésico	Leer y hacer	Activa o interacción
Imagen representativa	Imagen o significado que el estudiante considera que mejor representa a la palabra o significante en lengua extranjera	Icónica o concreta	Visual	Ver	Pasiva o receptiva
Pronunciación	Expresión hablada de la palabra o significante en lengua extranjera, la pronunciación sería un significante hablado	Apoyo auditivo	Auditivo	Escuchar	Pasiva o receptiva

En la Figura 5 se muestra un concepto general de relación sobre el material de estudio del método propuesto, en el cual: se da lectura a un texto, se agregan palabras desconocidas en lengua extranjera al vocabulario de estudio, por último, se repasan las palabras desconocidas mediante una imagen representativa y la pronunciación de la palabra desconocida en lengua extranjera.



*Figura 5 Concepto general de relación del material de estudio del método propuesto*

### 4.1.2 Diseño de componentes

El método propuesto tuvo como principal objetivo la creación de un sistema de cómputo de aprendizaje de idiomas (llamado *sistema SAIC*) que implemente sus principios de aprendizaje y enseñanza. Esto significó la construcción de modelos de usuarios, material de estudio y componentes de interacción; además se diseñaron dos algoritmos que personalizan el material didáctico disponible en el sistema mediante la recomendación de textos y la repetición espaciada. La Figura 6 muestra un diagrama general de la interacción de componentes del sistema SAIC del método propuesto.

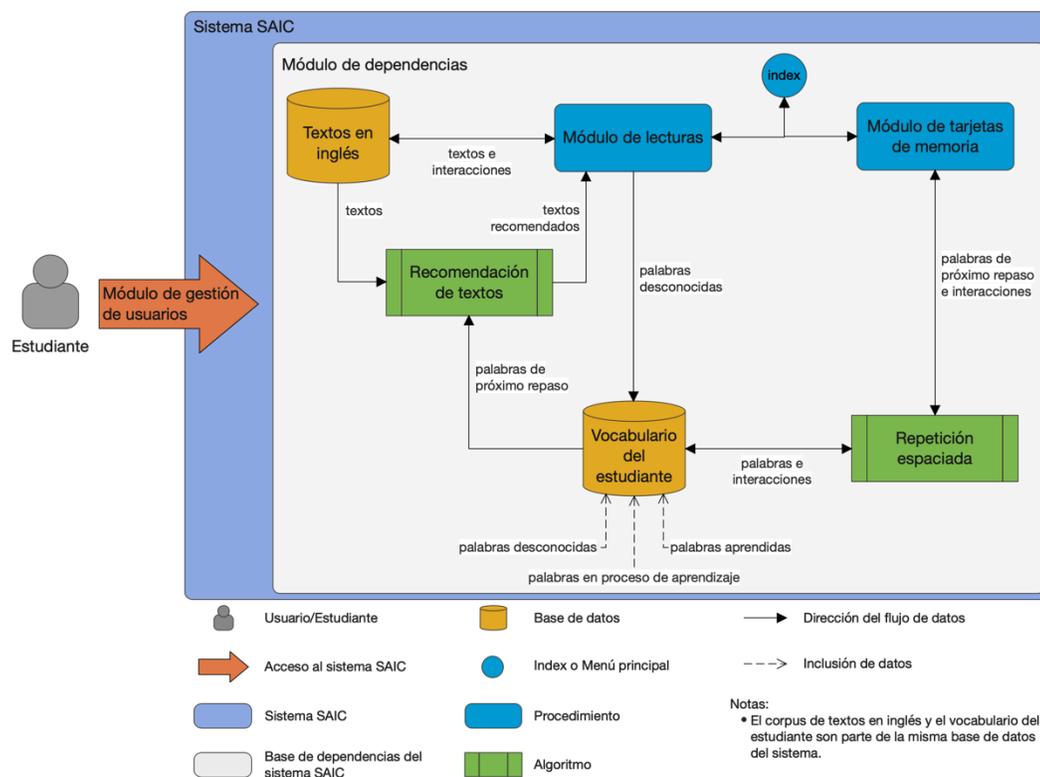


Figura 6 Interacción de componentes del sistema SAIC del método propuesto

#### 4.1.2.1 Modelo del estudiante

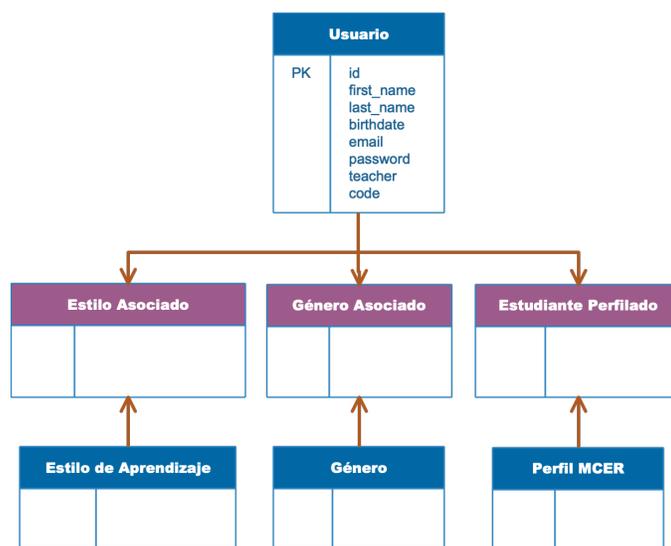
El sistema SAIC fue diseñado para estudiantes con un perfil mínimo deseable: haber concluido satisfactoriamente la Educación Básica de acuerdo con lo estipulado por la Secretaría de Educación Pública (2021) en México, tener conocimientos básicos sobre el uso de una computadora, conocer su perfilado en los estilos de aprendizaje del modelo VAK y un nivel de dominio del idioma inglés A1 hasta B2 de acuerdo con el MCER.

El *MCER* o *Marco común europeo de referencia para las lenguas* es una normativa de referencia que describe el nivel de dominio de los estudiantes de una lengua extranjera, este marco de referencia es ampliamente aceptado en Europa y cada vez más común en el resto del mundo (EF Education First, 2020). El *MCER* propone tres niveles de dominio divididos en dos subniveles cada uno para determinar las competencias lingüísticas de los estudiantes de una lengua extranjera (ver Tabla 8). En este trabajo, se recomienda utilizar la evaluación en línea del sitio Web *Cambridge Assessment English* (2020) para que los estudiantes conozcan su nivel de dominio en el idioma inglés.

*Tabla 8 Niveles de dominio de los estudiantes de una lengua extranjera del MCER (Consejo de Europa, 2001)*

Nivel	Subnivel	Descripción
A Básico	A1 Acceso	Entiende y utiliza expresiones de uso inmediato: dar información básica de su persona. Puede comunicarse si su interlocutor habla despacio y coopera
	A2 Plataforma	Es capaz de entender expresiones frecuentes, sabe comunicarse en tareas simples y cotidianas de forma breve (aspectos sociales)
B Independiente	B1 Intermedio	Comprende los puntos principales de textos en lengua estándar (sin tecnicismos), puede interactuar socialmente sin detallar o profundizar en temas específicos
	B2 Intermedio alto	Comprende textos técnicos cuando son de su especialización, no tiene problemas para el uso social y puede producir textos claros
C Competente	C1 Dominio operativo	Puede expresarse con buena fluidez, puede producir textos claros y estructurados, hace uso del idioma para fines sociales o profesionales
	C2 Maestría	Comprende todo lo que lee y escucha, reconstruye y resume información. Es capaz de describir temas específicos a detalle y con fluidez

A continuación, la Figura 7 expone las asociaciones de los componentes lógicos *Usuario*, *Perfil MCER*, *Estilo de Aprendizaje* y *Género* que componen el modelo del estudiante.



*Figura 7 Diagrama simplificado del modelo del estudiante*

#### 4.1.2.2 Modelo del material de estudio principal

El sistema SAIC contempla el material didáctico principal con el siguiente perfil: textos cortos de 300 a 500 palabras cada uno (1 cuartilla aproximadamente), pueden ser de género abierto (cuentos, historias, anécdotas o noticias), estar escritos en idioma inglés y etiquetados por nivel de complejidad de acuerdo con el MCER.

El *MCER*, como normativa de referencia, también describe el nivel de complejidad del material de estudio de una lengua extranjera, con el objetivo de asignar el material adecuado a cada estudiante. Los niveles de complejidad del material de estudio corresponden a los niveles de dominio de los estudiantes de una lengua extranjera (mostrados anteriormente en la Tabla 8). Con el objetivo de clasificar el material de estudio principal, en este trabajo se realizó una investigación sobre herramientas de clasificación del nivel de complejidad de textos de acuerdo con la normativa del MCER. La Tabla 9 muestra las herramientas de software analizadas.

*Tabla 9 Comparativa de herramientas de software para clasificar textos por su complejidad de acuerdo con el MCER*

Herramienta	Propietario	Disponibilidad	Método de clasificación
Text Inspector	WebLingua	13,000 palabras por 59.99 euros	No informado
Vocab Kitchen	Jeremy Garner	Gratuito	El algoritmo califica la dificultad de un texto comparando sus palabras con la <i>General Service List</i> de 1953. Esta lista cuenta con 2,000 palabras de uso frecuente del idioma inglés
Text Analyzer	RoadToGrammar	Gratuito	El algoritmo califica la dificultad de un texto comparando sus palabras con una lista de las 10,000 palabras más usadas en inglés, teniendo en cuenta su posición en la lista. Además, toma en cuenta el promedio de la longitud de las palabras y de las oraciones del texto

La herramienta *Text Inspector* (Weblingua Ltd, 2021) no informa el método de clasificación que utiliza y tiene un costo de suscripción de 59.99 euros al mes (aproximadamente 1,200 pesos mexicanos); por otro lado, *Vocab Kitchen* (Garner, 2021) utiliza una clasificación de las 2,000 palabras más utilizadas del idioma inglés (provenientes de la *General Service List*) para determinar qué tan común es el vocabulario y promediar un nivel de complejidad del texto; por último, *Text Analyzer* (RoadToGrammar, 2019) cuenta con un algoritmo de clasificación de complejidad de textos con base en una lista de 10,000 palabras, además de analizar la longitud de oraciones y palabras. Con base en el análisis anterior, se selecciona y recomienda el uso de *Text Analyzer* para obtener el nivel de complejidad de un texto como material de estudio de una lengua extranjera.

RoadToGrammar (2019) describe la herramienta *Text Analyzer* como un analizador de texto capaz de calcular el nivel de complejidad de un texto de acuerdo con el MCER, esta herramienta es utilizada por docentes de inglés para determinar el nivel aproximado de competencia para el cual es adecuado un texto y para obtener una lista de vocabulario sugerido. De acuerdo con su desarrollador, *Text Analyzer* calcula el nivel de complejidad de un texto mediante dos métricas: índice de complejidad del lenguaje, se obtiene comparando las palabras del texto de entrada contra una lista de las 10,000 palabras más usadas del inglés para obtener la posición de cada palabra en dicha lista; e índice *Flesch-Kincaid*, este índice analiza principalmente la longitud de las oraciones y la longitud de las palabras del texto de entrada. RoadToGrammar (2019) informa que la lista de 10,000 palabras más usadas del inglés es obtenida del análisis de un corpus de 200,000 palabras.

A continuación, la Figura 8 muestra las asociaciones de los componentes lógicos *Texto*, *Perfil MCER* y *Palabra de Texto* que componen el modelo del material de estudio principal, así como la relación con un componente lógico perteneciente al modelo del estudiante que establece un atributo de pertenencia entre un texto y un estudiante.

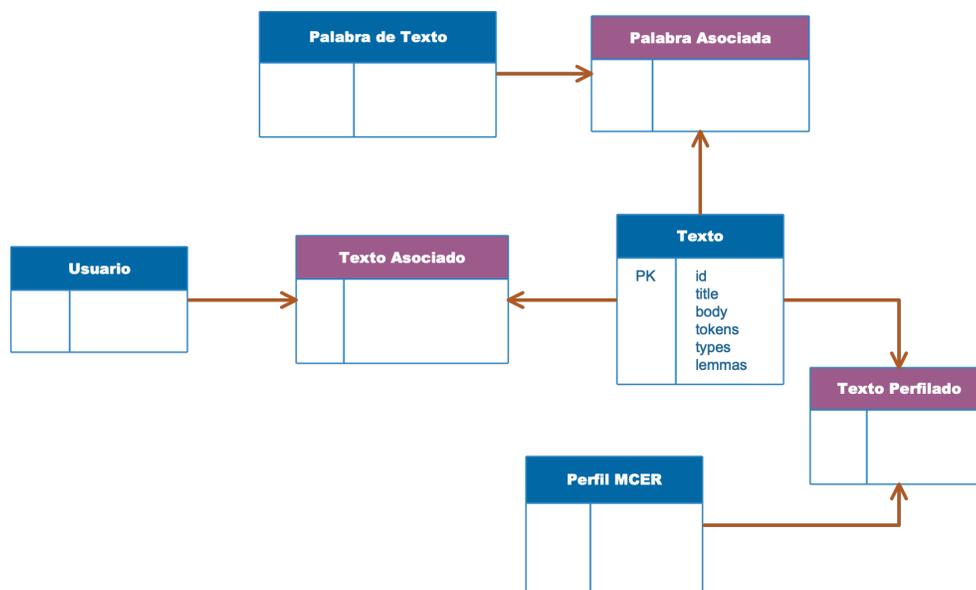


Figura 8 Diagrama simplificado del modelo del material de estudio principal

#### 4.1.2.3 Modelo del vocabulario del estudiante

El vocabulario del estudiante está constituido por un conjunto de palabras en lengua extranjera, este conjunto abarca palabras desconocidas, en proceso de aprendizaje y aprendidas.

A nivel computacional, el vocabulario es una interrelación entre la palabra, el usuario y el texto; esta interrelación se encarga de almacenar los atributos en común y establecer una pertenencia única entre los componentes, además de proporcionar un medio sencillo para realizar consultas de información. La Figura 9 presenta la interrelación *Vocabulario* formada por los componentes lógicos *Palabra de Usuario*, *Usuario* y *Texto*.

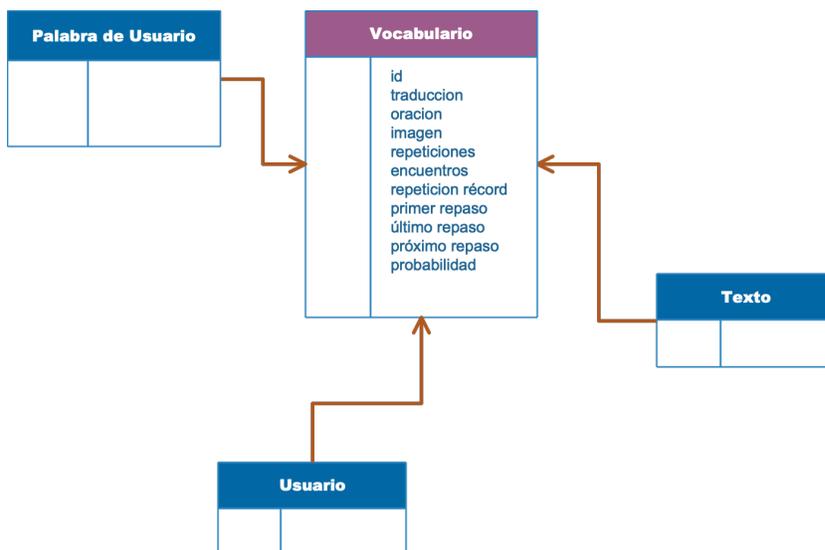


Figura 9 Diagrama simplificado del modelo del vocabulario del estudiante

#### 4.1.2.4 Modelo del material de estudio secundario

El material didáctico secundario corresponde a las tarjetas de memoria, en las que cada tarjeta contiene un elemento del *Vocabulario*. Este elemento cuenta con los siguientes atributos visibles en una tarjeta de memoria: imagen representativa de la palabra en lengua extranjera, pronunciación de la palabra en lengua extranjera, oración de contexto (oración en la que se encontró la palabra), traducción a su lengua materna y número de letras que componen la palabra. El objetivo consiste en que el estudiante logre recordar y asociar la palabra en lengua extranjera con su pronunciación y su imagen representativa. La Figura 10 muestra un prototipo donde se pueden observar los atributos visibles de una tarjeta de memoria, su número de tarjeta y un campo de texto donde el estudiante puede escribir la palabra que está repasando.

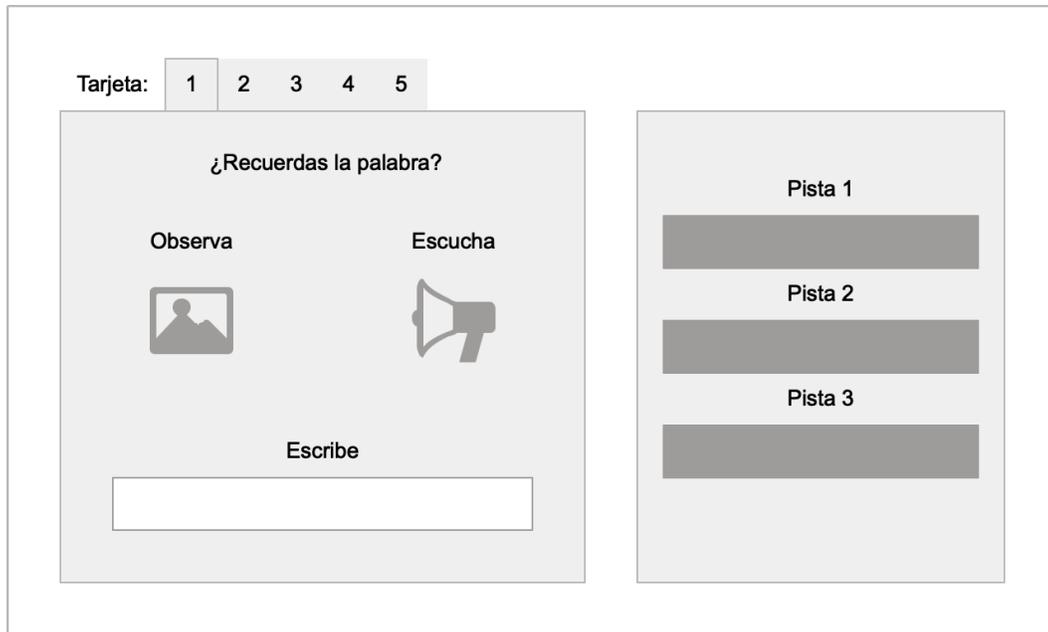


Figura 10 Prototipo de tarjeta de memoria

#### 4.1.2.5 Modelo de datos

El modelo de datos es una representación abstracta de la información del sistema SAIC y su interrelación. Se ha utilizado el *modelo relacional de bases de datos* para diseñar este modelo debido a su fundamento en la lógica de predicados y la teoría de conjuntos, estos fundamentos también forman parte de los principales sistemas manejadores de bases de datos relacionales; por lo tanto, este tipo de modelado garantiza la compatibilidad y transparencia del diseño del modelo con la base de datos implementada. La Tabla 10 presenta las primitivas utilizadas del modelo relacional y los atributos personalizados.

Tabla 10 Primitivas del modelo relacional y atributos personalizados

Nombre	Descripción	Primitiva		
Entidad	Representa un objeto o concepto del mundo real	<div style="background-color: #0070C0; color: white; padding: 2px; text-align: center; font-weight: bold;">Entidad</div> <table border="1" style="width: 100%; height: 30px; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%;"></td> </tr> </table>		
Atributo	Representa una característica que identifica a una entidad	Atributo1 Atributo2		
Relación	Representa una interacción entre dos entidades	<div style="background-color: #800080; color: white; padding: 2px; text-align: center; font-weight: bold;">Relación</div> <table border="1" style="width: 100%; height: 30px; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%;"></td> </tr> </table>		

Nombre	Descripción	Primitiva
Cardinalidad	Representa el grado de participación de una entidad en una relación	 Una A se asocia con una B
		 Una A se asocia con ninguna o una B
		 Una A se asocia con una o muchas B
		 Una A se asocia con ninguna o muchas B
Llave primaria	Atributo(s) designado(s) para identificar inequívocamente a la entidad. Por lo tanto, no se repite y siempre debe tener un valor	PK <i>Primary Key</i>
Llave foránea	Atributo designado para referenciar instancias de una entidad con otra entidad	FK <i>Foreign Key</i>
Atributo nulo	Personalizado. Si un nombre de atributo tiene este color, significa que puede ser nulo ( <i>null</i> )	Atributo
Atributo único	Personalizado. Si un nombre de atributo tiene este color, significa que debe ser un valor único ( <i>unique</i> )	Atributo
Grupo único	Personalizado. Si varios nombres de atributos tienen este color, significa que en su conjunto deben ser únicos ( <i>unique group</i> )	Atributo

El modelo de datos se diseñó con nombres de entidades, atributos y relaciones en el idioma inglés debido a la compatibilidad con los sistemas manejadores de bases de datos y la simplicidad en su nomenclatura. De acuerdo con lo anterior, una entidad tiene un nombre en singular del componente u objeto del mundo real que representa, un atributo tiene un nombre correspondiente a la característica del objeto representado y una relación tiene un nombre en singular de una entidad que asocia más un adjetivo que denota la relación entre las entidades involucradas.

Cabe señalar que este modelo de datos se encuentra bajo la tercera forma normal (3NF) dentro del concepto de *normalización* descrito por Lucas y Meseguera (2016) como la simplificación y organización coherente de los datos, con el objetivo de mantener su atomicidad, minimizar la redundancia, asegurar la consistencia e independencia de la información.

La Figura 11 presenta el modelo de datos del sistema SAIC, este modelo es la unión del modelo del estudiante, el modelo del material de estudio principal y el modelo del vocabulario del estudiante.

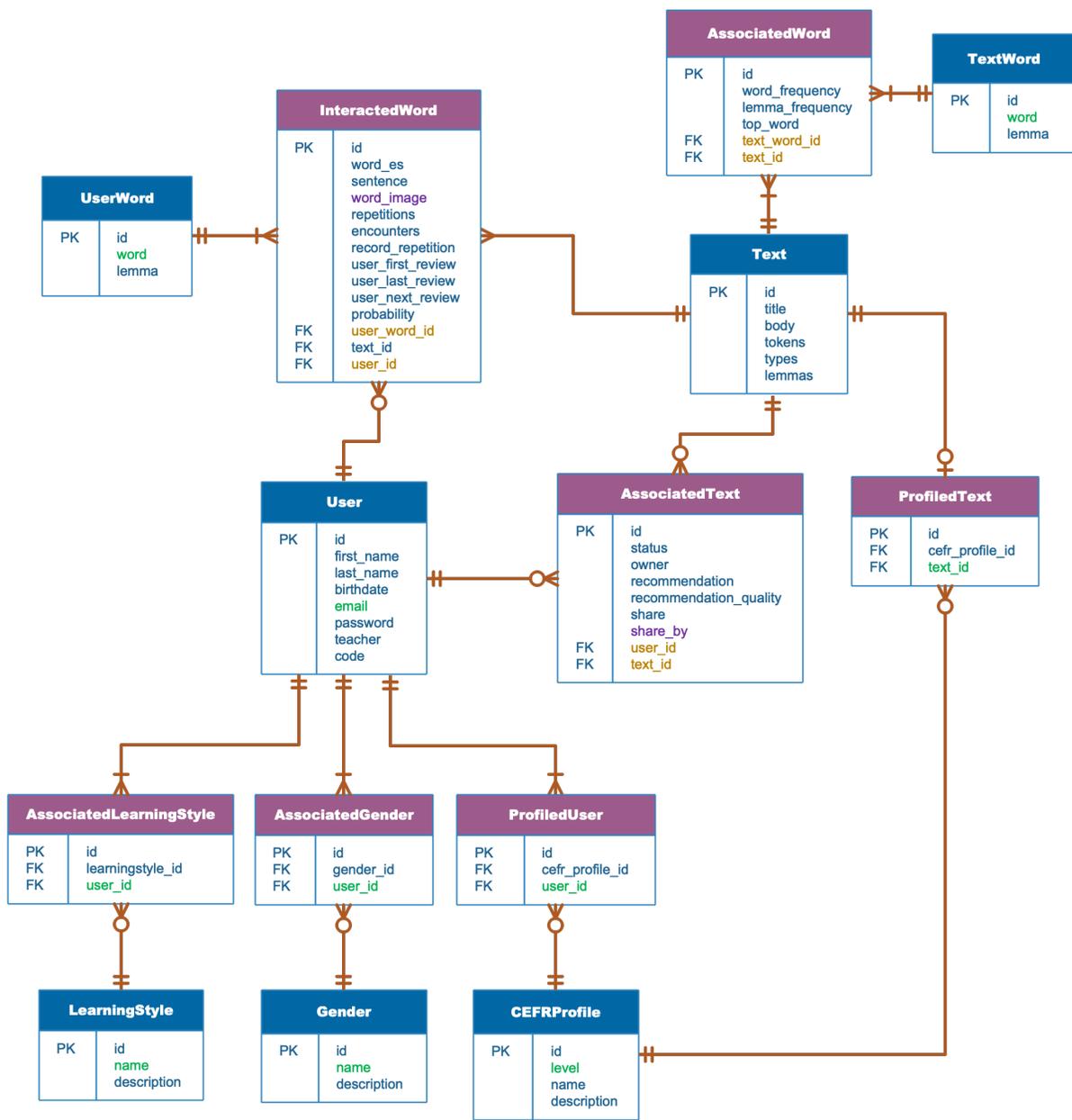


Figura 11 Modelo de datos del sistema SAIC

#### 4.1.2.6 Diseño del algoritmo de repetición espaciada

La *repetición espaciada* es una técnica de memorización que permite aprender información mediante repasos en intervalos de tiempo, estos intervalos se hacen más grandes a medida que el estudiante va memorizando la información. Esta técnica puede ser implementada por medio de distintos algoritmos o teorías, en este trabajo se utiliza la teoría de la curva del olvido de Ebbinghaus (1885) como base para calcular el momento de repaso de cierto material didáctico.

*La teoría de la curva del olvido* es descrita por Chmiel y Schubert (2018) como un mapeo sobre el decremento a través del tiempo de la capacidad de retención de la información de reciente adquisición. De acuerdo con la literatura, existen distintas ecuaciones que responden a esta teoría, incluyendo la ecuación original de Ebbinghaus (1885), sin embargo, en este trabajo se utilizó la ecuación (1) propuesta por Hu y Ogihara (2011) que tiene fundamento matemático en el decremento exponencial y permite un acceso sencillo a cada una de las variables.

$$f(t, s) = e^{-\frac{t}{s}} \quad (1)$$

En (1),  $f(t, s)$  es la retención de la memoria o la probabilidad de recordar cierta información de estudio, sus valores pueden ser  $[0, 1]$ ;  $t$  es el tiempo desde el último repaso (o intervalo de tiempo entre repasos), expresado en días; y  $s$  es el número de repasos que se han realizado, expresado en enteros positivos. La ecuación (1) se utiliza para calcular la retención de la memoria, no obstante, se puede despejar  $t$  para calcular el intervalo de tiempo entre repasos con la finalidad de obtener la fecha y hora del próximo repaso de cierto material de estudio. A continuación, se observa la manera de despejar  $t$  en la ecuación (1).

$$f(t, s) = e^{-\frac{t}{s}} \quad (1)$$

$$f(t, s) = p \text{ (} p \text{ de probabilidad, ya no será una función de } t, s \text{)}$$

$\therefore$

$$p = e^{-\frac{t}{s}}$$

$$\ln p = \ln e^{-\frac{t}{s}} \text{ (usando logaritmos)}$$

$$\ln p = -\frac{t}{s} \cdot \ln e \text{ (usando propiedades de los logaritmos)}$$

$$\ln p = -\frac{t}{s} \cdot 1 \text{ (logaritmo natural de } e = 1 \text{)}$$

$$s \cdot \ln p = -t \text{ (despejar } t \text{)}$$

$$(-1) \cdot s \cdot \ln p = (-1) \cdot -t \text{ (despejar } t \text{ en término positivo)}$$

$\therefore$

$$t = -s \cdot \ln p \quad (2)$$

En (2),  $t$  es el tiempo desde el último repaso (o intervalo de tiempo entre repasos), expresado en días;  $s$  es el número de repasos que se han realizado expresado en enteros positivos; y  $p$  es la retención de la memoria o la probabilidad de recordar cierta información de estudio, sus valores pueden ser  $[0, 1]$ .

De acuerdo con la literatura, en este trabajo se asume un mínimo de 50% de *probabilidad de recordar cierta información de estudio*, es decir, se establece un mínimo de  $p = 0.5$  con la finalidad de lograr aprender la información deseada. Además, el *número de repasos* ( $s$ ) se incrementa en 1 si el repaso que realice el estudiante es correcto y consecutivo, en otro caso, toma un valor de 0. La Tabla 11 muestra el intervalo de tiempo de los diez primeros repasos espaciados calculados con la ecuación (2).

*Tabla 11 Diez primeros repasos espaciados calculados*

Número de repasos ( $s$ )	Retención de la memoria ( $p$ )	$-s$	$\ln p$	Intervalo de tiempo entre repasos (días) ( $t$ )
1	0.5	-1	-0.6931	0.6931
2	0.5	-2	-0.6931	1.3862
3	0.5	-3	-0.6931	2.0794
4	0.5	-4	-0.6931	2.7725
5	0.5	-5	-0.6931	3.4657
6	0.5	-6	-0.6931	4.1588
7	0.5	-7	-0.6931	4.8520
8	0.5	-8	-0.6931	5.5451
9	0.5	-9	-0.6931	6.2383
10	0.5	-10	-0.6931	6.9314

El *intervalo de tiempo entre repasos* ( $t$ ) proporciona la posibilidad de obtener la fecha y hora del próximo repaso de cierto material de estudio a partir de la fecha y hora de su último repaso, esto se puede representar matemáticamente con la ecuación (3).

$$\text{próximo repaso} = \text{último repaso} + t \quad (3)$$

La Tabla 12 presenta los diez próximos repasos espaciados de cierta información de estudio con una probabilidad del 50% de recordarla, también se puede observar que los intervalos de tiempo entre repasos se hacen mayores con cada iteración correcta y consecutiva.

*Tabla 12 Diez próximos repasos espaciados calculados*

Último repaso	Número de repasos (s)	Intervalo de tiempo entre repasos (días) (t)	Intervalo de tiempo entre repasos (horas) (t)	Próximo repaso
2019-12-01 15:00	1	0.6931	≈17	2019-12-02 08:00
2019-12-02 08:00	2	1.3862	≈33	2019-12-03 17:00
2019-12-03 17:00	3	2.0794	≈50	2019-12-05 19:00
2019-12-05 19:00	4	2.7725	≈66	2019-12-07 21:00
2019-12-07 21:00	5	3.4657	≈83	2019-12-11 08:00
2019-12-11 08:00	6	4.1588	≈99	2019-12-15 11:00
2019-12-15 11:00	7	4.8520	≈116	2019-12-20 07:00
2019-12-20 07:00	8	5.5451	≈133	2019-12-25 20:00
2019-12-25 20:00	9	6.2383	≈149	2020-01-01 01:00
2020-01-01 01:00	10	6.9314	≈166	2020-01-07 23:00

Con base en lo anterior, se ha diseñado un *algoritmo de repetición espaciada* que permite darle seguimiento al estudio del material didáctico, con la finalidad de sugerir un próximo repaso y clasificar dicho material de acuerdo con su complejidad para el estudiante. Este algoritmo se debe ejecutar inmediatamente después de que el estudiante realice un repaso de una tarjeta de memoria (material de estudio secundario). A continuación, se muestran los pasos del algoritmo de repetición espaciada propuesto.

- Paso 1: suponiendo que el algoritmo para calcular el próximo repaso de una tarjeta de memoria se implementa en un método llamado `calculate_next_review()`, este método debe recibir como parámetros: el elemento del vocabulario del estudiante (`InteractedWord`) que se ha utilizado para la tarjeta de memoria recién repasada y el resultado del repaso (si el estudiante recordó la palabra en lengua extranjera, se establece como `CORRECT_REVIEW`; en caso contrario `WRONG_REVIEW`).
- Paso 2: si el resultado del repaso es que el estudiante recordó la palabra en lengua extranjera (`CORRECT_REVIEW`), entonces el número de repeticiones (`repetitions`) y la repetición récord (`record_repetition`) del elemento del vocabulario del estudiante se incrementa en 1; en caso contrario (`WRONG_REVIEW`), el número de repeticiones (`repetitions`) del elemento del vocabulario del estudiante se asigna a 1.
- Paso 3: el número de encuentros (`encounters`) con ese elemento del vocabulario del estudiante se incrementa en 1.

- Paso 4: se calcula el intervalo de tiempo entre repasos (*time\_range* o *t*) mediante la ecuación (1) con una probabilidad mínima de recordar la información de estudio del 50% (*MINIMUM\_PROBABILITY*) y se actualiza esta probabilidad utilizada en el elemento del vocabulario del estudiante (*probability*).
- Paso 5: se obtiene la fecha actual y se asigna como la fecha del último repaso (*user\_last\_review*) del elemento del vocabulario del estudiante.
- Paso 6: se calcula la fecha del próximo repaso (*user\_next\_review*) del elemento del vocabulario del estudiante mediante la expresión (3).
- Paso 7: el método *calculate\_next\_review()* devuelve el elemento del vocabulario del estudiante actualizado.

La Figura 12 presenta un pseudocódigo que implementa el algoritmo de repetición espaciada descrito en los pasos anteriores.

```

Paso 1 calculate_next_review(interacted_word, review_note)
Paso 2   if review_note == CORRECT_REVIEW then
           interacted_word.repetitions++
           interacted_word.record_repetition++
         else // en caso contrario: WRONG_REVIEW
           interacted_word.repetitions = 1
         end
Paso 3   interacted_word.encounters++
Paso 4   time_range = - interacted_word.repetitions * Math.log_n(MINIMUM_PROBABILITY)
           interacted_word.probability = MINIMUM_PROBABILITY
Paso 5   interacted_word.user_last_review = DateTime.now()
Paso 6   interacted_word.user_next_review = interacted_word.user_last_review + time_range
Paso 7   <- interacted_word
         end

```

*Figura 12 Pseudocódigo que implementa el algoritmo de repetición espaciada*

### 2.1.2.7 Diseño del algoritmo de recomendación de textos

Un *sistema de recomendación* tiene como objetivo combatir la sobrecarga de información que existe en los sistemas de cómputo. Este objetivo se puede lograr mediante el análisis, filtrado y recuperación de información relevante de acuerdo con determinado perfil, interacción o preferencia.

En este trabajo, se propone un algoritmo de recomendación de textos que permite al estudiante reencontrar las palabras de próximo repaso de su vocabulario de estudio en textos

distintos al texto de origen, además, encontrar una palabra en distintos contextos propicia que el estudiante pueda ubicar nuevas palabras relacionadas para agregarlas a su vocabulario. Este algoritmo se diseñó con base en cuatro tópicos: base de información, métrica TF-IDF, espacio vectorial y similitud coseno.

La *base de información* se refiere a los datos de entrada del algoritmo de recomendación que están compuestos por los textos disponibles en el sistema SAIC y el vocabulario del estudiante. Este tipo de datos de entrada no puede ser analizado por un método matemático, sin embargo, existen métricas que se utilizan para crear representaciones numéricas de los textos. En este contexto, la *métrica TF-IDF (Term Frequency - Inverse Document Frequency)* mide la relevancia que tiene una palabra de un texto en comparación con toda la colección de textos a la que pertenece, es decir, el valor de una palabra usando TF-IDF aumenta proporcionalmente al número de veces que aparece en su texto, pero es compensada inversamente proporcional al número de textos en los que aparece dicha palabra dentro de la colección de textos; esto permite ponderar las palabras más raras (Allahyari, et al., 2017). Se puede obtener el peso TF-IDF de una palabra mediante la ecuación (4).

$$\text{Peso TF IDF} = TF \cdot IDF \quad (4)$$

En donde:

*Term Frequency (TF)* mide la frecuencia de aparición de la palabra dentro del texto, esta medida se obtiene con la ecuación (5).

$$TF = \frac{\text{Número de veces que la palabra aparece en el texto}}{\text{Número total de palabras en el texto}} \quad (5)$$

*Inverse Document Frequency (IDF)* mide la importancia de la palabra dentro de la colección de textos, es decir, determina si la palabra aparece en muchos textos o en pocos (si es común o rara); esta medida se obtiene con la ecuación (6).

$$IDF = \frac{\text{Número total de textos en la colección}}{\text{Número de textos en los que aparece la palabra}} \quad (6)$$

La vectorización de un texto mediante la métrica TF-IDF proporciona un peso de relevancia (valor escalar) por cada palabra del texto, es decir, se obtiene un *vector* de pesos de relevancia TF-IDF que representan al texto; esto convierte a los textos disponibles del sistema SAIC en un *espacio vectorial*. Este tipo de representación permite indexar, recuperar, clasificar o comparar

textos por medio de sus vectores, en este caso, se requiere comparar dos vectores para determinar su similitud semántica. La *similitud coseno* es una métrica que permite determinar la similitud entre dos vectores no vacíos mediante el coseno del ángulo de dichos vectores (Allahyari, et al., 2017). La similitud coseno se obtiene con cualquier variante de la ecuación (7).

$$\text{Similitud coseno} = \cos \theta = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (7)$$

En (7),  $\theta$  es el ángulo entre los vectores,  $A$  y  $B$  son los vectores,  $A_i$  y  $B_i$  son los elementos o atributos de los vectores. Esta ecuación tiene tres resultados principales: *similitud coseno* = 1, los vectores son exactamente iguales; *similitud coseno* = 0, los vectores son ortogonales; y *similitud coseno* = -1, los vectores son opuestos. De acuerdo con lo anterior, si la similitud coseno se acerca a 1, los textos son más similares semánticamente; mientras que, si la similitud coseno se acerca a 0, los textos son poco o nada similares semánticamente; por último, cabe señalar que la métrica TD-IDF no genera pesos negativos, por lo tanto, los resultados para la similitud coseno de este análisis sólo podrá tener valores en el rango [0, 1]. La Figura 13 muestra una representación trigonométrica de la similitud coseno de dos vectores para los escenarios válidos.

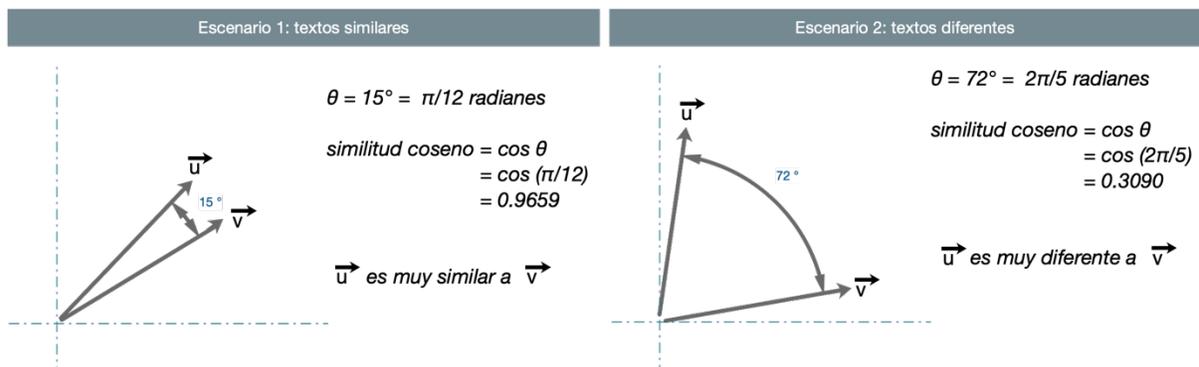


Figura 13 Representación trigonométrica de la similitud coseno para vectores TF-IDF

Con base en lo anterior, se ha diseñado un *algoritmo de recomendación de textos* que implementa las tres fases de los sistemas de recomendación en conjunto con técnicas de procesamiento de lenguaje natural (PLN), con el objetivo de crear representaciones vectoriales de los textos del sistema SAIC, analizar dichas representaciones y proporcionar una recomendación al estudiante de acuerdo con su vocabulario de próximo repaso.

El algoritmo de recomendación de textos emplea una técnica de filtrado basado en contenido, es decir, se filtran los textos del sistema SAIC de acuerdo con el análisis de su contenido y atributos; y se utiliza una retroalimentación implícita, en otras palabras, se analiza la interacción que tiene el estudiante con su vocabulario; con el objetivo de comprar los textos y la retroalimentación para poder realizar una recomendación de textos al estudiante. De acuerdo con las capacidades que posea el servidor que dé soporte al sistema SAIC, se sugiere condicionar la actualización de textos recomendados, con la finalidad de no sobrecargar el servidor e impedir una reducción en su rendimiento. A continuación, se expone el algoritmo de recomendación de textos propuesto.

- *Preprocesamiento.* Se aplican las siguientes técnicas de PLN a los textos antes de que formen parte del sistema SAIC: *normalización*, se remueven elementos no textuales y se sustituyen las contracciones del idioma inglés por palabras completas, así como los números por sus nombres; *lematización*, se convierten las palabras a su forma base mediante el modo *host* de *FreeLing*; y *tokenización*, se divide el texto en palabras y se omiten los signos de puntuación, un *token* es una palabra.
- *Actividad de exploración.* La actividad consiste en que el estudiante identifique las palabras que conoce de un conjunto de diez palabras en lengua extranjera y retroalimente su vocabulario con las palabras desconocidas. Esta actividad es obligatoria y se presenta inmediatamente después del primer inicio de sesión en el sistema SAIC, su principal objetivo es amortiguar el *problema del inicio en frío* de este tipo de algoritmos.
- *Fase de recolección de información.* Se recolecta la información desde la base de datos del sistema SAIC: primero, se obtienen los textos que el estudiante no haya leído y que posean un nivel de complejidad igual o mayor a su perfil MCER; después, se obtiene el vocabulario de cada texto del paso anterior y hasta veinte palabras de próximo repaso del vocabulario del estudiante; por último, se remueven las palabras de próximo repaso de la lista de *stopwords* y se omiten las palabras de la nueva lista de *stopwords* del vocabulario de cada texto.
- *Fase de aprendizaje.* Se utiliza la medida TF-IDF para vectorizar el vocabulario de cada texto y las palabras de próximo repaso del vocabulario del estudiante, se genera un vector de cada texto y uno más de dichas palabras. Después, se compara cada vector de los textos

con el vector de dichas palabras mediante la similitud coseno, con el objetivo de obtener la similitud semántica entre cada texto y las palabras de próximo repaso.

- *Fase de predicción.* El algoritmo realiza una predicción directa, es decir, selecciona diez textos que hayan tenido la mayor similitud semántica con las palabras de próximo repaso del vocabulario del estudiante y los proporciona como una recomendación de textos.

La Figura 14 muestra la interacción entre los procedimientos de las fases del algoritmo de recomendación de textos que ha sido descrito anteriormente.

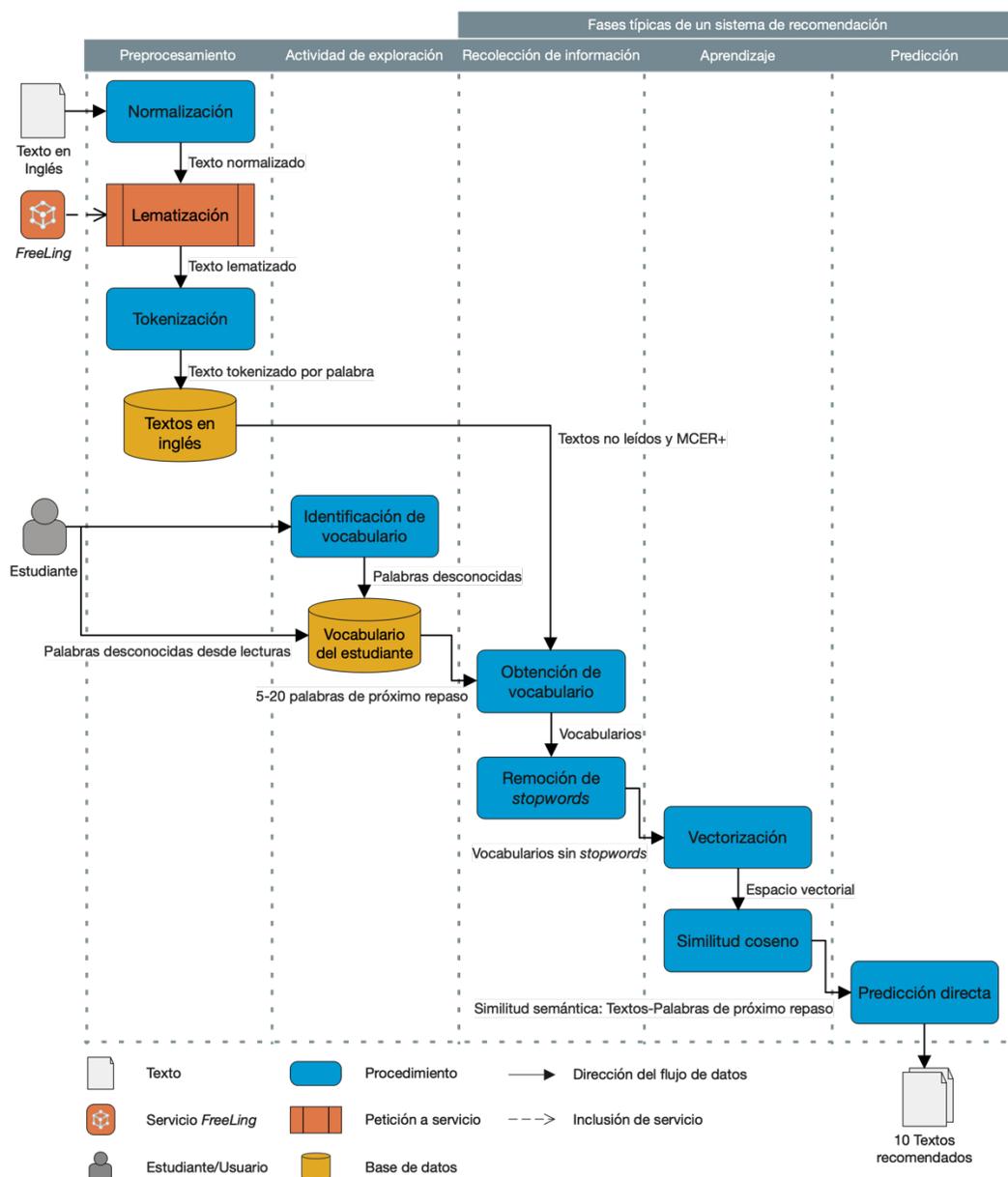


Figura 14 Interacción entre procedimientos de las fases del algoritmo de recomendación de textos

### 4.1.3 Actividades de aprendizaje

Las actividades de aprendizaje describen una visión general sobre la forma en que el estudiante adquiere conocimientos mediante el sistema SAIC, además, las actividades son propuestas tomando en cuenta los principios de aprendizaje y enseñanza de este método. Estas actividades se relacionan de tal manera que generan una retroalimentación de información, con la finalidad de personalizar el material didáctico para cada estudiante. La Figura 15 muestra la interacción que mantienen las actividades con otros mecanismos del sistema SAIC.

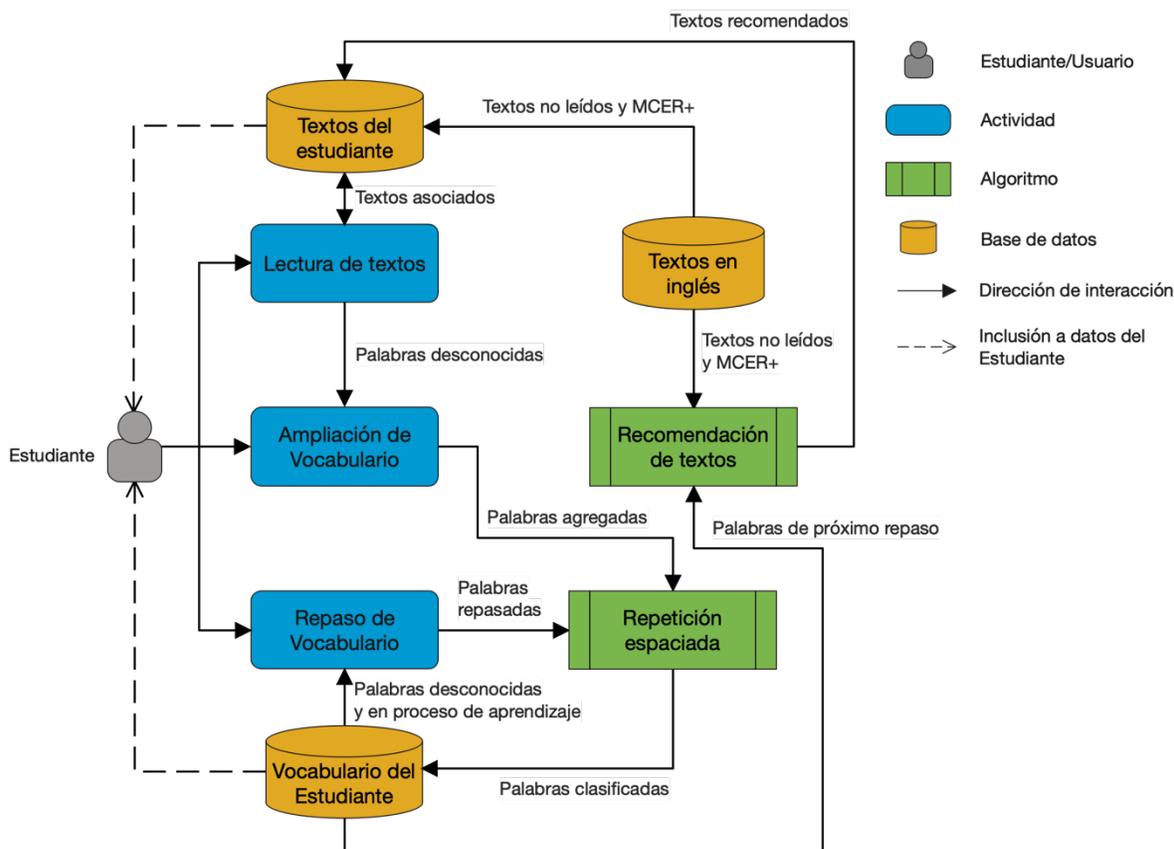


Figura 15 Interacción de las actividades del estudiante con otros componentes del sistema SAIC

#### 4.1.3.1 Lectura de textos

El estudiante practica su comprensión sobre una lengua extranjera mediante la lectura de textos, también le permite encontrar palabras desconocidas en lengua extranjera. Además, esta actividad cuenta con el algoritmo de recomendación de textos que personaliza una lista de textos de acuerdo con el vocabulario del estudiante. El estudiante cuenta con las siguientes funciones que apoyan su aprendizaje en esta actividad:

- *Listado de lecturas disponibles* de acuerdo con el nivel de dominio de lengua extranjera del estudiante, además de filtros de lecturas por características: nuevas, leídas, recomendadas y subidas por el estudiante.
- *Recomendación de textos* de acuerdo con el vocabulario de próximo repaso del estudiante.
- *Traducción de palabras desconocidas* encontradas por el estudiante durante la lectura de un texto.

#### **4.1.3.2 Ampliación del vocabulario**

El estudiante puede agregar o guardar palabras desconocidas en lengua extranjera que encuentre durante una lectura; además, tiene la posibilidad de relacionar dicha palabra con una imagen representativa, con la finalidad de tener claro el significado de la palabra y que le ayude a recordarla durante sus repasos. El estudiante tiene las siguientes funciones para ampliar su vocabulario:

- *Marcado de la palabra desconocida en contexto* durante la lectura permite saber qué palabra está seleccionada y lista para agregarse al vocabulario del estudiante.
- *Pronunciación de la palabra marcada* cuando se intenta agregar al vocabulario del estudiante.
- *Traducción de la palabra marcada* cuando se intenta agregar al vocabulario del estudiante.
- *Asociación de la palabra marcada con una imagen representativa* por medio de una galería de imágenes sugeridas.

#### **4.1.3.3 Repaso del vocabulario**

El estudiante tiene la posibilidad de repasar las palabras desconocidas que ha agregado a su vocabulario durante sus lecturas mediante tarjetas de memoria, con la finalidad de aprender y reforzar dichas palabras. Esta actividad cuenta con seguimiento de aprendizaje de las palabras mediante el algoritmo de repetición espaciada que sugiere el próximo repaso y clasifica las palabras de acuerdo con su dificultad para el estudiante; además, esta información de

seguimiento de aprendizaje también es utilizada por el algoritmo de recomendación de textos. El estudiante cuenta con las siguientes funciones para realizar su repaso de vocabulario:

- *Modelo de tarjeta de memoria con imagen representativa y pronunciación*, además de oración de contexto, traducción y cantidad de letras de la palabra desconocida como apoyo de repaso.
- *Listado de tarjetas de memoria* de acuerdo con la complejidad de las palabras desconocidas para el estudiante, se organizan en los filtros de tarjetas: próximas, caducadas, en espera y desafío.
- *Repaso con 5 o 10 tarjetas de memoria* que el estudiante puede seleccionar de acuerdo con su tiempo disponible y el número de palabras que desea repasar en un momento determinado.

## 4.2 Etapa de desarrollo

### 4.2.1 Especificaciones técnicas del sistema SAIC

El sistema SAIC está basado en la Web, en otras palabras, es un sistema que requirió el uso de tecnologías Web para su construcción, así mismo, requiere dichas tecnologías para su mantenimiento, extensión y ejecución. Las especificaciones técnicas brindan la información básica sobre las tecnologías Web y configuraciones que necesita un desarrollador para conformar el entorno de desarrollo y ejecución que requiere el sistema SAIC. En la Tabla 13 se pueden observar las especificaciones sobre las que se debe montar un entorno de desarrollo.

*Tabla 13 Especificaciones técnicas del sistema SAIC*

Especificación	Selección	Descripción
Plataforma	Web	El sistema SAIC se ha desarrollado con base en tecnologías Web con el objetivo de facilitar el acceso a los estudiantes, esto debido a que sólo es necesario que tengan un navegador Web actualizado en su computadora y acceso a Internet.
Sistema operativo	MacOS 10.15 Linux5+	Los comandos y componentes se han documentado para MacOS 10.15, sin embargo, muchos comandos son compatibles con sistemas con núcleo Linux 5 o UNIX. Si estos comandos o componentes no son compatibles, se debe seguir la referencia de la fuente oficial de la tecnología para tener acceso a las instrucciones para el sistema operativo adecuado.

Especificación	Selección	Descripción
Base de datos	PostgreSQL 13.1	El sistema manejador de bases de datos PostgreSQL (SMBD) es proporcionado por <i>The PostgreSQL Global Development Group</i> (2020). Este sistema manejador dispone de una instalación gratuita y transparente con el modelo relacional, se ajusta a la CPU y memoria disponible (escalabilidad), cumple con las características <i>ACID Compliant</i> (garantiza la atomicidad, consistencia, aislamiento y durabilidad de los datos), amplía sus funciones mediante extensiones (extensibilidad) e implementa el estándar SQL (estándar ISO/IEC 9075:2011).
Lenguaje de programación	Ruby	Es un lenguaje de programación orientado a objetos, potente para la Web con su gema <i>rails</i> , compatible con PostgreSQL por medio de su gema <i>pg</i> y robusto frente a las tareas de procesamiento de lenguaje natural.
Control de versiones	Bitbucket	El servicio Bitbucket proporciona hospedaje de control de versiones Git para evitar incidentes como pérdida de información e incompatibilidad de cambios durante el desarrollo de software.

Después de cumplir con las especificaciones anteriores, se deben configurar las siguientes dependencias mediante la línea de comandos, los comandos deben ejecutarse desde el directorio raíz del sistema SAIC.

- *Ruby on Rails*. Se compone de la instalación del intérprete del lenguaje de programación Ruby mediante el software RVM (*Ruby Version Manager*) y de la gema *rails* (gema es el nombre que recibe una librería en el lenguaje Ruby) proporcionada por Papis, Kuczynski y Seguin (2020).

Instalación de *GPG keys*:

```
gpg2 --recv-keys 409B6B1796C275462A1703113804BB82D39DC0E3
7D2BAF1CF37B13E2069D6956105BD0E739499BDB
```

Instalación de RVM, interprete del lenguaje Ruby y gema *rails*:

```
\curl -sSL https://get.rvm.io | bash -s stable --rails
```

- *Configuración para SMBD PostgreSQL*. El sistema SAIC se configuró para el SMBD PostgreSQL, sin embargo, podría emplearse otro sistema manejador si se realiza la configuración adecuada. En este caso, se requiere la gema *pg* para que el sistema SAIC pueda comunicarse con el SMBD PostgreSQL.

Instalación de la gema *pg*:

```
gem install pg
```

En el sistema SAIC, la configuración de conexión con el SMBD PostgreSQL se encuentra en el archivo ubicado en la ruta `/config/database.yml`.

- *Gestión de paquetes y configuración de estilo.* Se debe configurar un gestor de paquetes que facilita la instalación de componentes de estilo.

Instalación del gestor de paquetes *webpacker* de la gema *rails*:

```
bundle exec rails webpacker:install
```

Instalación de la librería de estilos de interfaz Web *bootstrap*:

```
yarn add bootstrap
```

Instalación de *jquery*, dependencia de *bootstrap*:

```
yarn add jquery
```

Instalación de *pepper.js*, dependencia de *bootstrap*:

```
yarn add pepper.js
```

En el sistema SAIC, el archivo de estilos en formato *css/sass* se puede encontrar en la ruta `/app/javascript/css/style.scss`. Después, se importa *bootstrap* y el archivo de estilos en el documento con la ruta `/app/javascript/packs/application.js` como se observa en la Figura 16.

```
11 // KHS: importar bootstrap despues de agregarlo - yarn add bootstrap
12 import 'bootstrap'
13 // KHS: crear carpeta y archivo, importarlo
14 import 'css/styles'
```

*Figura 16 Configuración de librerías de estilos*

Por último, se compilan los cambios con el gestor *webpacker* por medio de la instrucción:

```
./bin/webpacker-dev-server
```

- *Instalación de gemas mediante el archivo Gemfile.* El entorno *Ruby on Rails* proporciona un archivo de configuración para agregar e instalar gemas de manera sencilla. El desarrollador sólo requiere agregar el nombre de la gema y su versión al archivo de la ruta `/Gemfile` y ejecutar el siguiente comando para instalar la gema indicada:

```
bundle install
```

Por ejemplo, si se desea instalar la gema *pagy* para paginar una lista de lecturas o tarjetas de memoria; primero, se debe agregar el nombre de la gema y su versión como se muestra en la Figura 17.

```
27 # KHS: Paginación
28 gem 'pagy', '~> 3.8' # omit patch digit and use the latest if possible
```

*Figura 17 Configuración del archivo Gemfile para la gema pagy*

Después, se debe ejecutar el comando:

```
bundle install
```

- *Configuración del gestor de usuarios.* El lenguaje de programación Ruby posee la gema *devise* que proporciona la seguridad necesaria para administrar cuentas de usuarios y sus sesiones. Esta gema se puede instalar mediante el archivo `/Gemfile`, agregando el nombre de la gema y su versión como se muestra en la Figura 18.

```
25 # KHS: Manejo de sesiones
26 gem 'devise'
```

*Figura 18 Configuración del archivo Gemfile para la gema devise*

Después, se instala la gema *devise* mediante el comando:

```
bundle install
```

Por último, se instalan los componentes de la gema *devise* para el sistema SAIC con el comando:

```
rails generate devise:install
```

- *Archivos de configuración relevantes del sistema SAIC.* Estos archivos contienen distintas configuraciones que pueden ser relevantes para el mantenimiento, extensión o ejecución del sistema SAIC, por lo tanto, se enlistan para que se tenga un acceso rápido a estas configuraciones.

Archivo de configuración de rutas:

```
/config/routes.rb
```

Archivo de configuración de idiomas:

```
/config/locales/es.yml
```

```
/config/locales.devise.es.yml
```

Archivo de configuración de estilos:

```
/app/javascript/css/style.scss
```

Archivo de configuración de base de datos:

```
/config/database.yml
```

Ruta de los archivos de la estructura de la base de datos, se recomienda crear la base de datos mediante las herramientas proporcionadas por el entorno *Ruby on Rails*:

```
/db/
```

Archivo de configuración de dependencias:

```
/Gemfile
```

- *Ejecución local del sistema SAIC*. Una vez configurado el entorno de desarrollo, se puede ejecutar el sistema SAIC en modo local (es decir, en la misma computadora) mediante la función *server* de la gema *rails*.

Comando para la ejecución local:

```
rails server
```

Este comando provee el sistema SAIC como un servicio Web, esto significa que se puede tener acceso local al sistema con un navegador Web mediante una de las siguientes direcciones Web:

```
http://localhost:3000
```

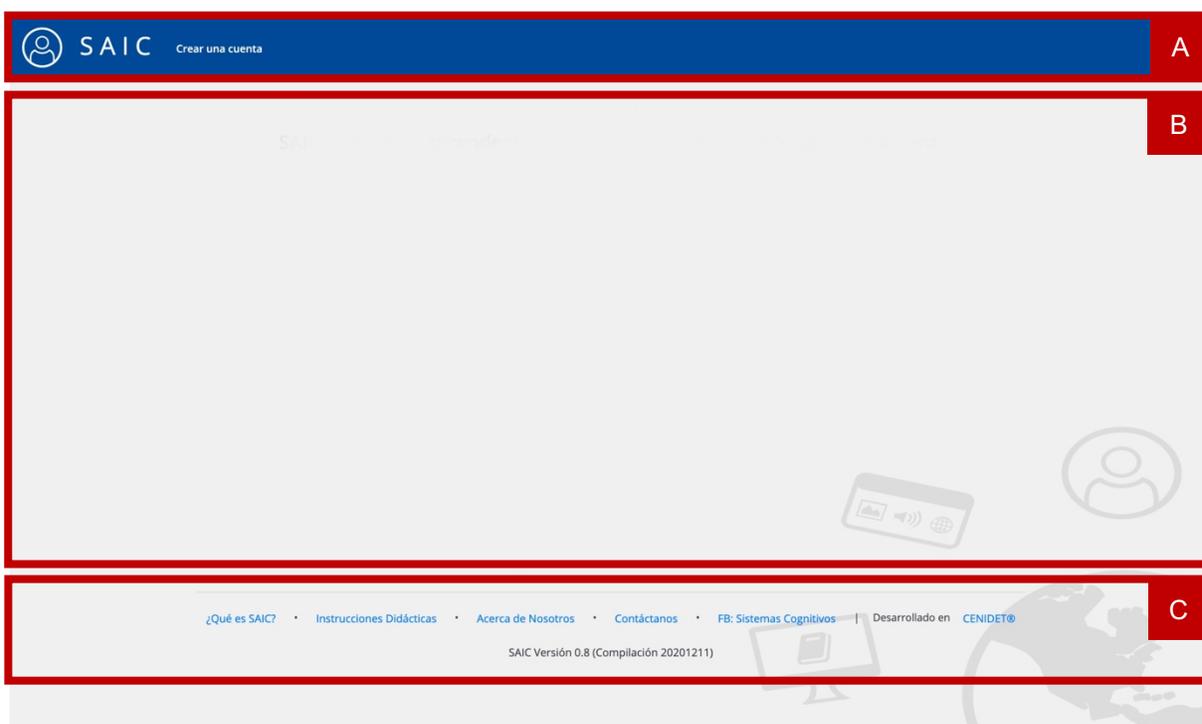
```
http://127.0.0.1
```

#### ***4.2.2 Interfaz del sistema SAIC***

La RAE (2021) define la interfaz de un sistema de cómputo como un componente físico o lógico que posibilita la comunicación entre el usuario y la computadora. La interfaz del sistema SAIC es un componente lógico que permite al estudiante realizar las *actividades de aprendizaje* mediante un concepto enfocado en el contenido y colores usados de manera uniforme en todo el sistema.

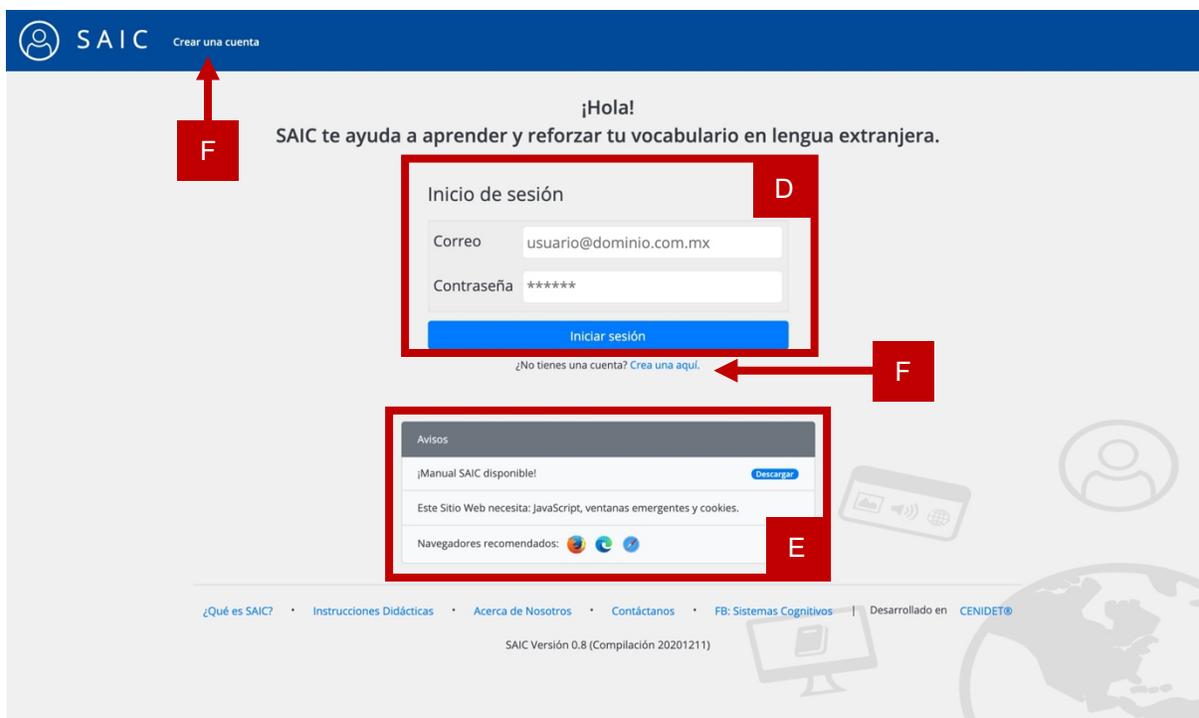
El acceso a través de Internet del sistema SAIC se ha configurado por medio de un servidor que proporciona una dirección IP pública y provee el sistema como servicio Web. Una vez configurado, se puede acceder a la interfaz del sistema desde un navegador Web moderno como Firefox, Safari o Edge; sólo se requiere registrarse en el sistema e iniciar sesión con los datos registrados. El sistema SAIC tiene una interfaz compuesta por distintas páginas y vistas, sin embargo, se han seleccionado las 12 vistas más relevantes para presentarlas a continuación.

La *plantilla* es una vista base que está presente en todas las páginas que conforman al sistema, con la finalidad que el estudiante reconozca que se encuentra dentro del sistema, así como proporcionarle accesos rápidos a la información, ayuda y opciones relevantes. Esta plantilla se compone de tres partes: *encabezado* (A), contiene el logo del sistema y accesos rápidos a funciones de acuerdo con el contexto de la página; *cuerpo* (B), muestra el contenido principal de la página; y *pie* (C), presenta enlaces a información y ayuda relevante del sistema, así como a sitios de interés. La Figura 19 muestra la plantilla de las páginas Web del sistema SAIC.



*Figura 19* Plantilla del sistema SAIC

La *página de inicio de sesión* da la bienvenida a los estudiantes al sistema SAIC que no tienen una cuenta o que no han iniciado sesión en el sistema. Esta página se compone de dos secciones: *inicio de sesión* (D), permite entrar al sistema a través de un correo y contraseña previamente registrado; y *avisos* (E), muestra mensajes notificando de las novedades del sistema a los estudiantes; también cuenta con enlaces para registrar una cuenta nueva para estudiantes que no cuenten con una (F). La Figura 20 presenta la página de inicio de sesión.



*Figura 20* Página de inicio de sesión del sistema SAIC

La *página de registro de cuenta* permite registrar los datos de un estudiante en el sistema SAIC, con la finalidad que pueda tener acceso con su correo electrónico y contraseña registrada a dicho sistema. Esta página se compone de un *formulario* que solicita los siguientes datos: correo electrónico, nombre, apellido, fecha de nacimiento, género, estilo de aprendizaje, perfil MCER, contraseña y confirmación de la contraseña. Todos los campos son obligatorios y cuentan con ayuda para poder llenarlos; además, el *campo fecha de nacimiento* (G) cuenta con un calendario nativo del navegador Web para llenar este campo (el diseño y disposición de este calendario puede variar de acuerdo con el navegador Web utilizado), el *campo estilo de aprendizaje* (H) cuenta con el *cuestionario de estilos de aprendizaje del modelo VAK* (Secretaría de Educación Pública, 2004) a modo de un *test* para que el estudiante pueda conocer su estilo y el *campo perfil MCER* (I) cuenta con un enlace a un *test* o evaluación en línea proporcionado por el sitio Web *Cambridge Assessment English* (2020) para que el estudiante conozca su perfil MCER. La Figura 21 muestra la página de registro de cuenta.

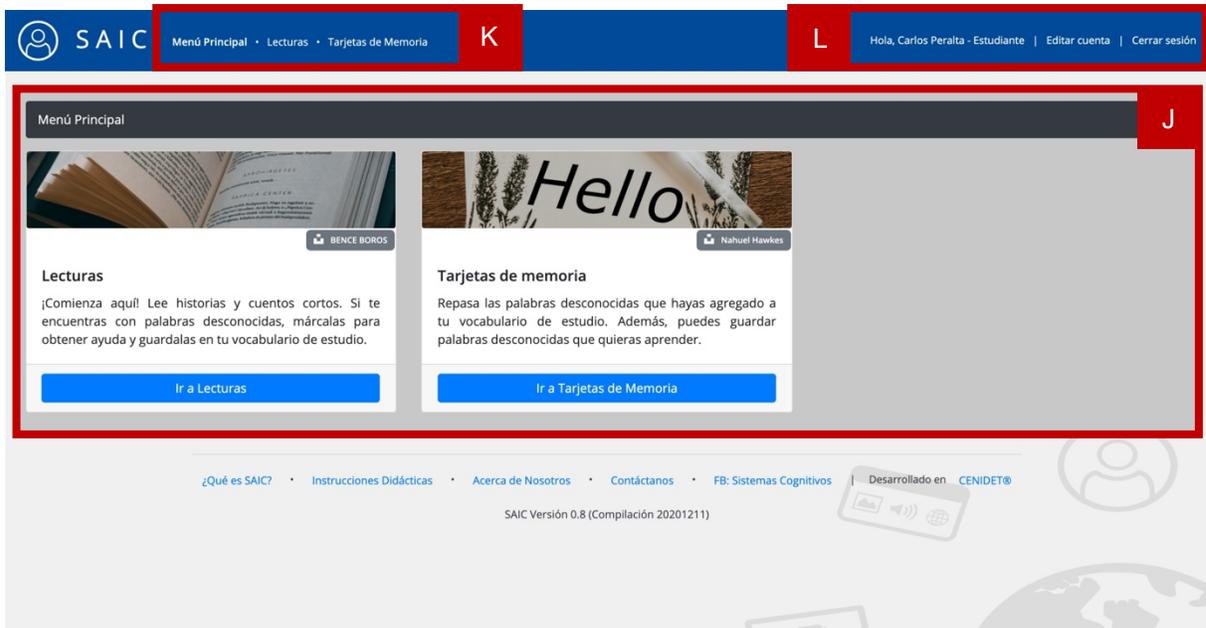
Creando una cuenta para tener acceso a la plataforma SAIC.

Registro de cuenta

Correo	Ej. usuario@dominio.com.mx
Nombre	Ej. Carlos
Apellido	Ej. Quintero
Fecha de Nacimiento	dd/mm/aaaa  <b>G</b>
Género	Selecciona una opción 
Estilo de Aprendizaje	Selecciona una opción  <b>H</b> <small>¿No sabes tu Estilo de Aprendizaje? Realiza un test aquí.</small>
Perfil MCER	Selecciona una opción  <b>I</b> <small>¿No sabes tu Perfil MCER? Realiza un test aquí.</small>
<small>Tu Contraseña debe cumplir con:</small> <ul style="list-style-type: none"> <li>- Mínimo 6 caracteres.</li> <li>- Máximo 128 caracteres.</li> <li>- Sólo letras y números.</li> </ul>	
Contraseña	*****
Confirmación	*****
<input type="button" value="Crear cuenta"/>	
<input type="button" value="Cancelar"/>	

*Figura 21* Página de registro de cuenta del sistema SAIC

La *página del menú principal* se presenta una vez que el estudiante ha iniciado sesión en el sistema SAIC. Esta página tiene tres secciones: *menú principal* (J), proporciona una imagen y descripción sobre la actividad de lectura de textos y repaso de tarjetas de memoria; *opciones del menú principal* (K), provee enlaces a las actividades del menú principal en el encabezado de la página; y *opciones de gestión de la cuenta del estudiante* (L), muestra el nombre del estudiante en el encabezado de la página, así como opciones para editar sus datos de registro y cerrar su sesión actual. La Figura 22 muestra la página del menú principal.



*Figura 22* Página del menú principal del sistema SAIC

La *vista de la actividad de exploración* consiste en un conjunto de diez palabras en lengua extranjera que el estudiante tiene que identificar como conocidas o desconocidas para retroalimentar su vocabulario de estudio, esto permite amortiguar el problema de inicio en frío del algoritmo de recomendación de textos. Esta vista se presenta inmediatamente después del inicio de sesión del estudiante en el sistema mientras no se haya realizado; una vez realizada la actividad de exploración, no se vuelve a mostrar al estudiante. La Figura 23 muestra la vista de la actividad de exploración.

**Prueba inicial**

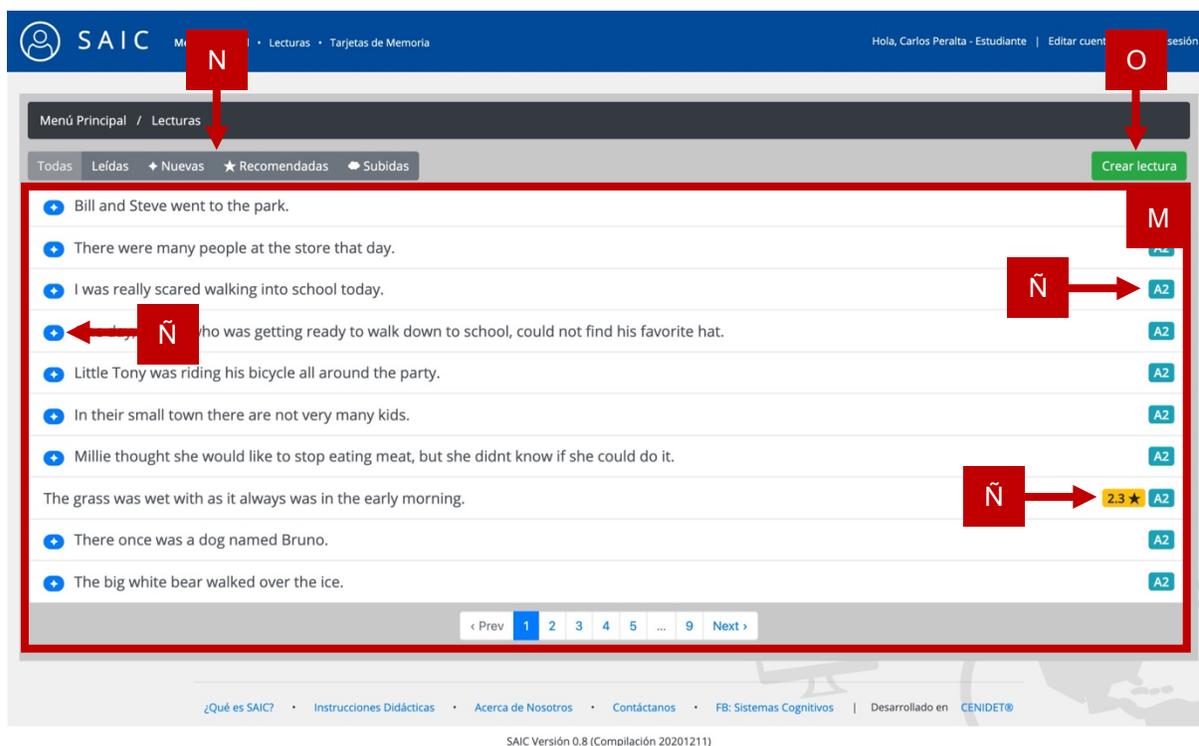
¡Hola! Necesitamos algunas palabras en inglés para tu vocabulario de estudio. Selecciona sólo las palabras que conozcas de la siguiente lista, cuando estés listo da clic en el botón **Aceptar**.

- afraid
- brushes
- woods
- flour
- playmates
- shares
- claws
- barked
- lettuce
- hungry

Aceptar
Cerrar sesión

*Figura 23* Vista de la actividad de exploración del sistema SAIC

La *página de la lista de lecturas* proporciona una lista paginada de lecturas disponibles para el estudiante (M), la paginación se realiza cada diez lecturas. Esta lista tiene los siguientes filtros de organización (N): todas las lecturas, lecturas leídas, lecturas recomendadas y lecturas subidas por el estudiante. También se utilizan las siguientes etiquetas (Ñ): lectura nueva, calificación de recomendación, perfil MCER y lectura subida por el estudiante. Además, se muestra un botón que permite al estudiante crear y subir una lectura al sistema (O). La Figura 24 presenta la página de la lista de lecturas.



*Figura 24* Página de la lista de lecturas del sistema SAIC

La *vista de lecturas recomendadas* muestra una lista de diez lecturas generadas por el algoritmo de recomendación de textos con base en el vocabulario de próximo repaso del estudiante. Esta vista muestra un cartel con información acerca de las recomendaciones y su calificación que va de 5 estrellas para una lectura altamente recomendada a 0 estrellas para una lectura no recomendada (P). La Figura 25 expone la vista de lecturas recomendadas que se activa cuando el estudiante da clic sobre el filtro *recomendadas* de la página de la lista de lecturas.

Menú Principal / Lecturas recomendadas

Todas Leídas + Nuevas ★ Recomendadas Subidas Crear lectura

**Nuestras recomendaciones.** P

Lista de lecturas recomendadas con base en tu vocabulario de estudio, una lectura marcada con 5 ★ es altamente recomendada mientras que una lectura con 0 ★ no es recomendada.

The grass was wet with as it always was in the early morning.	2.3 ★ A2
+ One day a man named John was walking down the block near the park when he came across a butterfly.	1.7 ★ A2
+ Once upon a time a driver was taking some meat to the store.	3.1 ★ A2
+ Years back in a kingdom far away there lived a family in the woods.	2.7 ★ A2
+ Che was very upset.	2.6 ★ A2
+ One day 3 friends went to a pool party that they had been excited about for a long time.	2.0 ★ A2
+ Susan wanted to have a birthday party.	2.0 ★ A2
+ Tim and Janey woke up earlier than normal.	1.9 ★ A2
+ John was in the third grade, and nine years old.	1.8 ★ A2
+ The dog was playing in the street outside.	1.8 ★ A2

< Prev 1 Next >

*Figura 25 Vista de lecturas recomendadas de la lista de lecturas del sistema SAIC*

La *página del visor de lectura* se compone de tres secciones: *opciones de la lectura* (Q), permiten volver a la lista de lecturas, así como editar o borrar la lectura sólo si el estudiante subió dicha lectura al sistema; *visor de lectura* (R), muestra el título de la lectura, su perfil MCER, su contenido y proporciona la posibilidad de traducir las palabras desconocidas dando clic sobre ellas; y *herramientas del visor* (S), ofrece opciones para guardar la palabra seleccionada en el vocabulario del estudiante, ocultar todas las traducciones visibles en el visor y desplazar el foco de la página a la parte superior de la lectura. La traducción de palabras es proporcionada por *Cloud Translation* de Google Inc. (2021). La Figura 26 muestra la página del visor de lectura.

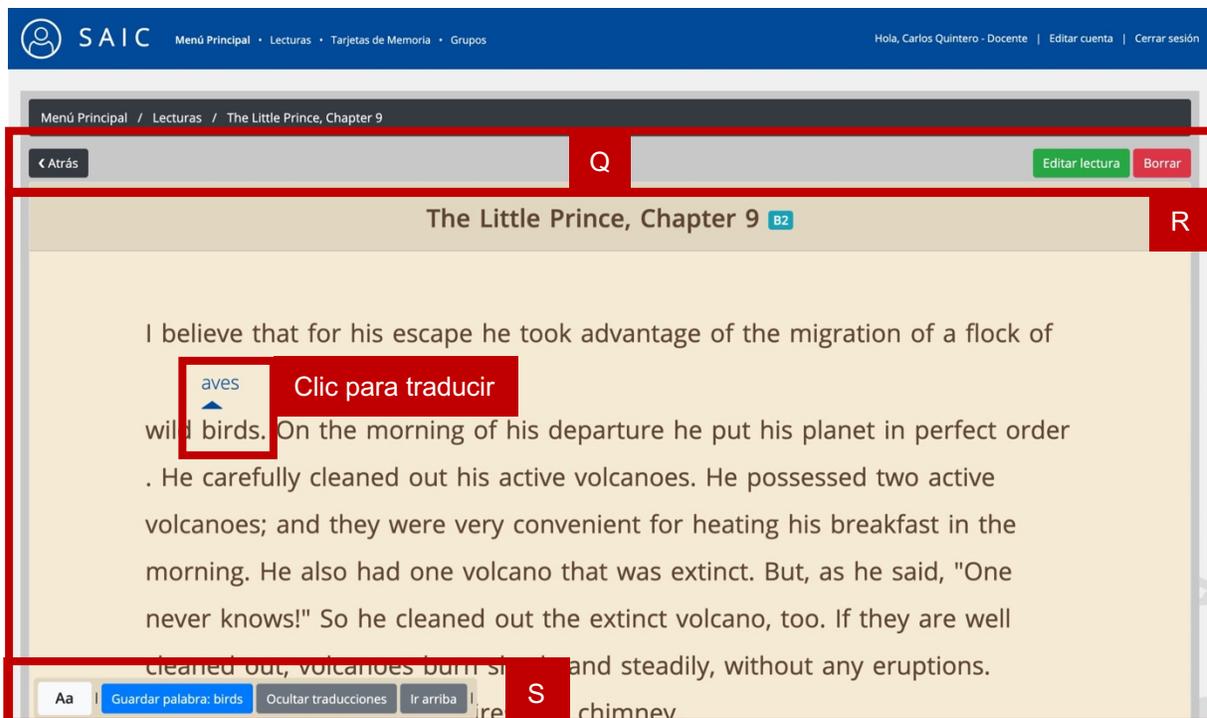


Figura 26 Página del visor de lectura del sistema SAIC

La *página de nueva lectura* permite crear o transcribir un texto para subirlo al sistema y que pueda estar disponible para todos los estudiantes de acuerdo con su perfil MCER. Esta página se muestra cuando el estudiante da clic en el botón *crear lectura* de la página de lista de lecturas. La página de nueva lectura se compone de cuatro secciones: *opciones de nueva lectura* (T), permite volver al visor de lectura, así como subir la lectura redactada; *requisitos de la nueva lectura* (U), proporciona información acerca de las características que deben tener los textos redactados; *campos de la lectura* (V), estos son campos de texto para introducir el título y contenido de la lectura; y *campo del perfil MCER* (W), provee un menú desplegable para seleccionar el perfil MCER de la lectura, además, proporciona un enlace hacia la herramienta *Text Analyzer* (RoadToGrammar, 2019) para obtener el perfil MCER de la lectura en caso que el estudiante lo desconozca. Todos los campos de la nueva lectura son obligatorios. La Figura 27 presenta la página de nueva lectura.

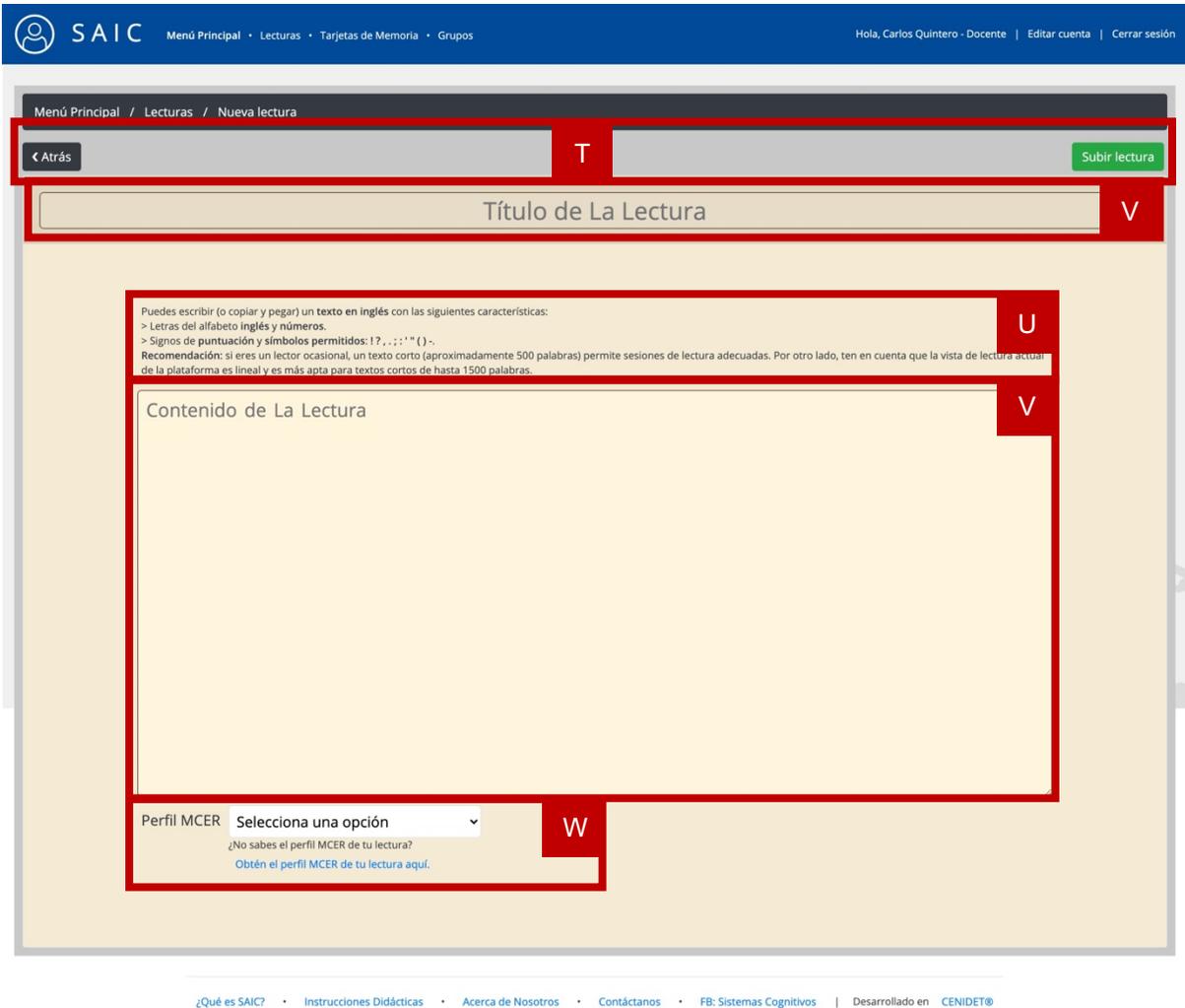


Figura 27 Página de nueva lectura del sistema SAIC

La *vista para guardar una palabra en el vocabulario de estudio* permite agregar la palabra seleccionada al vocabulario del estudiante para un futuro repaso con tarjetas de memoria. Esta vista se compone de cuatro secciones: *información de la vista* (X), presenta información sobre la importancia de ampliar el vocabulario de estudio; *palabra seleccionada* (Y), muestra un campo de texto no editable con la palabra que se va a guardar en el vocabulario del estudiante, así como un botón que permite reproducir la pronunciación de dicha palabra; *traducción de la palabra seleccionada* (Z), muestra un campo de texto con la traducción de la palabra seleccionada; y *galería de imágenes* (AA), presenta una galería de diez imágenes donde el estudiante puede elegir la imagen que mejor represente a la palabras seleccionada, con la finalidad de guardar la relación palabra-imagen en el vocabulario del estudiante. La traducción de palabras es proporcionada por *Cloud Translation* de Google Inc. (2021), la pronunciación de

palabras es proporcionada por *Cloud Text To Speech* de Google Inc. (2021) y la galería de imágenes es proporcionada por *Programmable Search Engine* de Google Inc. (2021). La Figura 28 muestra la vista para guardar palabra en tu vocabulario de estudio.

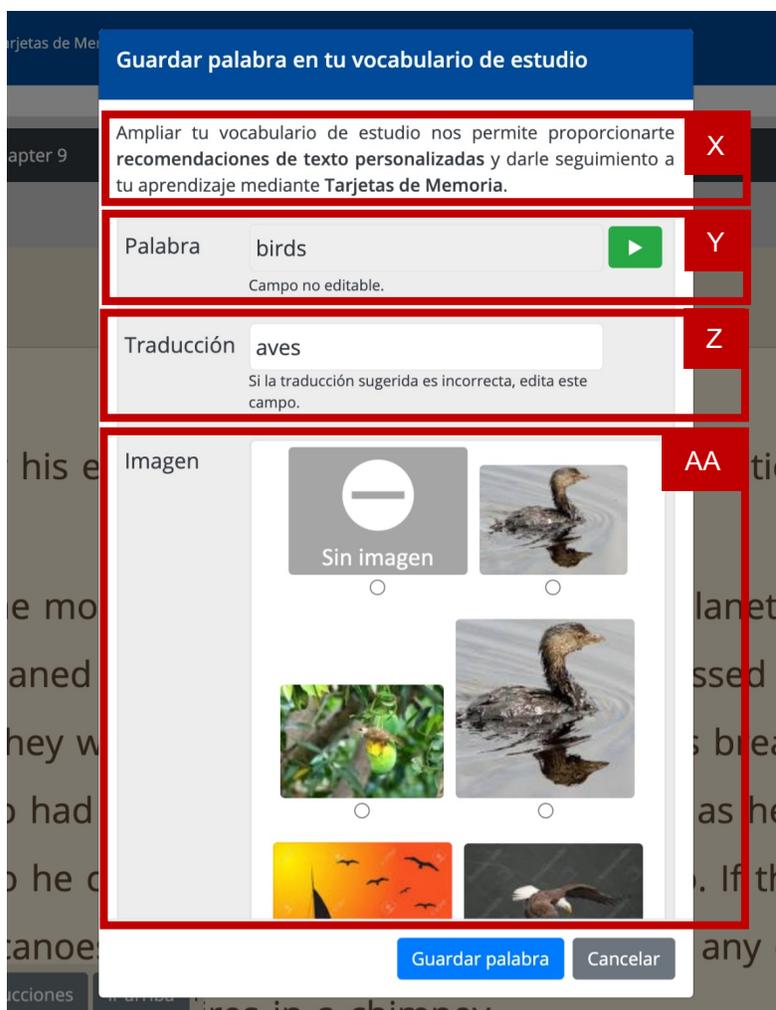
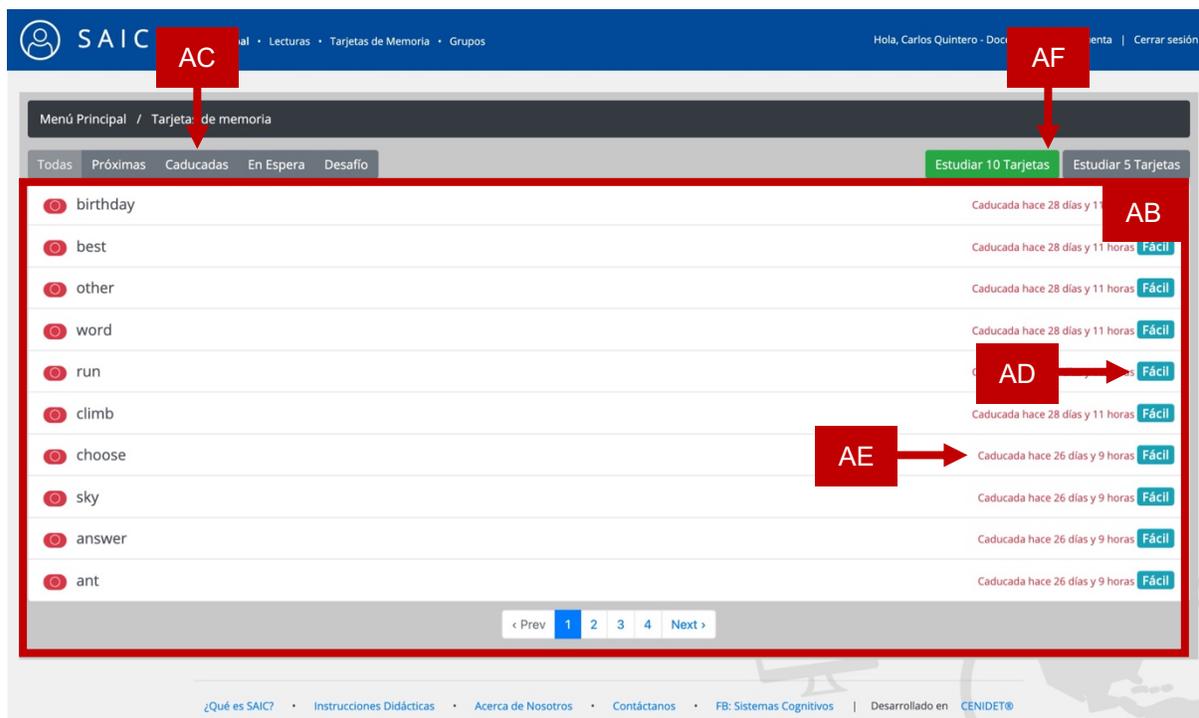


Figura 28 Vista para guardar una palabra en el vocabulario de estudio del visor de lectura del sistema SAIC

La página de la lista de tarjetas de memoria proporciona una lista paginada de palabras en proceso de aprendizaje del vocabulario del estudiante (AB), la paginación se realiza cada diez palabras. Esta lista tiene los siguientes filtros de organización (AC): todas las tarjetas, tarjetas próximas a estudiar, tarjetas caducadas, tarjetas en espera y tarjetas en desafío (contienen palabras especialmente difíciles para el estudiante). También se utiliza la etiqueta de complejidad de la palabra (AD) e información sobre la caducidad, espera y hora próxima de repaso de la tarjeta de memoria (AE). Además, se muestran dos botones que permiten al estudiante iniciar un repaso de cinco o diez tarjetas de memoria (AF) de acuerdo con su tiempo

disponible y el avance de estudio que desee realizar. La Figura 29 presenta la página de la lista de tarjetas de memoria.



*Figura 29* Página de la lista de tarjetas de memoria del sistema SAIC

La *página de repaso de vocabulario con tarjetas de memoria* presenta un botón que reproduce la pronunciación de la palabra en proceso de aprendizaje (AG) y su imagen relacionada (AH) que eligió el estudiante al momento de agregarla al vocabulario, con la finalidad que el estudiante recuerde y escriba la palabra en proceso de aprendizaje en el campo de texto disponible en la tarjeta (AI). El estudiante también se puede apoyar de la oración de contexto, su traducción y su número de letras de la palabra en proceso de aprendizaje para recordarla (AJ). Además, se muestra un cartel con instrucciones sobre el repaso de tarjetas de memoria (AK). Las características descritas se pueden observar en la Figura 30.

Menú Principal / Tarjetas de memoria / Repaso de vocabulario

Menú Principal / Tarjetas de memoria / Repaso de vocabulario

Repaso de palabras en Inglés. AK

Observa la imagen y escucha la pronunciación de la palabra desconocida. Después, escribe la palabra que recuerdes en el campo de texto.  
Nota: puedes apoyarte con la oración de contexto, la traducción al Español y el número de letras de la palabra.

Tarjeta 1 2 3 4 5

¿Recuerdas la palabra? AG

Observa AH

Escucha

Pronunciar

Escribe

Palabra en Inglés Escribe aquí la palabra que recuerdas AI

Siguiente Tarjeta

Ayuda AJ

Ayuda 1: Oración de Contexto

He'd give the queen shorts as \_\_\_\_\_ presents.

Ayuda 2: Traducción Al Español

cumpleaños

Ayuda 3: Número de Letras

8 Letras

¿Qué es SAIC? · Instrucciones Didácticas · Acerca de Nosotros · Contáctanos · FB: Sistemas Cognitivos | Desarrollado en CENIDET®

SAIC Versión 0.8 (Compilación 20201211)

Figura 30 Página de repaso de vocabulario con tarjetas de memoria del sistema SAIC

El sistema SAIC cuenta con un manual de usuario por función que amplía la información acerca de la interfaz presentada en esta sección.

## 4.3 Etapa de evaluación

### 4.3.1 Plan de implementación y evaluación

El plan de implementación y evaluación contempla una serie de actividades que se realizan antes, durante y después de que los estudiantes de una lengua extranjera han tenido acceso al sistema SAIC. Dentro de estas actividades, se da seguimiento al aprendizaje de los estudiantes durante el periodo de evaluación, con la finalidad de determinar si el uso del sistema SAIC aporta un mayor beneficio en su aprendizaje de vocabulario de una lengua extranjera en comparación con otro sistema que es su competencia más cercana.

Las *actividades previas* a que los estudiantes tengan acceso al sistema SAIC se dividen en dos tipos: 1) *Actividades del sistema*, consisten en la configuración de un servidor para proveer servicios Web con tecnologías *Ruby on Rails*, así como la exportación del código fuente del sistema al servidor para ponerlo en línea (accesible desde Internet), esto se ha realizado siguiendo la guía *Deploy Ruby On Rails* de GoRails LLC (2021). 2) *Actividades institucionales*, se realiza un convenio de participación con una institución educativa o con un docente de lenguas extranjeras, posteriormente, se selecciona un grupo de estudiantes para participar en la evaluación de acuerdo con el *modelo del estudiante* presentado en este documento, por último, se realiza una presentación del proyecto de tesis a los estudiantes participantes y se pone a su disposición un manual de usuario o presentación tutorial de uso del sistema.

Las *actividades de seguimiento* se llevan a cabo durante el periodo que los estudiantes están practicando con el sistema SAIC. Estas actividades consisten en *reuniones de control y reportes estadísticos* semanales sobre el progreso de los estudiantes en sus tareas dentro del sistema, con la finalidad de proveer evidencia a la institución educativa o al docente acerca de los datos recolectados, la forma de evaluación y el progreso de los estudiantes participantes por semana.

Las *actividades posteriores* a que los estudiantes han tenido acceso al sistema SAIC contemplan una *reunión de conclusión* y un *reporte final* que resume la información de aprendizaje recolectada por medio del sistema. Estas actividades tienen el objetivo de contrastar el progreso semanal del aprendizaje de los estudiantes durante el periodo de evaluación para determinar una tendencia. Con la finalidad de dar causalidad y veracidad al resultado obtenido de la evaluación del sistema SAIC, se realiza un análisis para conocer el beneficio en el aprendizaje de vocabulario de una lengua extranjera que proporciona el uso del sistema SAIC a los estudiantes comparado con el uso de otro sistema, además, se verifica el resultado obtenido del análisis anterior mediante una prueba de hipótesis.

#### ***4.3.2 Investigación experimental***

La investigación experimental es la estrategia que permite al investigador intervenir en las variables de un fenómeno, es decir, modificar unas variables (variables independientes o VI) que supone afectan a otras (variables dependientes o VD); con el objetivo de observar y medir el comportamiento del fenómeno para determinar los efectos causales del tratamiento o modificación aplicada (Bärnighausen, Röttingen, Rockers, Shemilt, & Tugwell, 2017). La

elaboración de un experimento conduce a la obtención de resultados que se utilizan comúnmente para validar o refutar una hipótesis de investigación.

Una *hipótesis* es una suposición que se toma como base en un razonamiento y su validez se determina mediante las conclusiones obtenidas de una investigación, esta validez es influenciada por la validez interna y externa. Por un lado, la *validez interna* se encarga de medir la veracidad de causalidad entre una variable independiente y una variable dependiente; por ejemplo, un experimento con un solo grupo no permite determinar si los resultados obtenidos son efecto del tratamiento o no, por el contrario, si se utiliza un grupo de control para contrastar los resultados, entonces es posible determinar los efectos del tratamiento de forma más precisa. Por otro lado, la *validez externa* se ocupa de medir si la investigación, en su totalidad, se puede generalizar con otra muestra, ambiente, tiempo o variaciones del tratamiento; por ejemplo, un experimento en el que no es posible o es difícil aplicar una asignación aleatoria se verá afectado con grupos no equivalentes que se pueden traducir en un sesgo en los resultados del experimento (Huitt, Hummel, & Kaeck, 1999).

Una investigación experimental o *diseño de un estudio* puede ser de tres tipos: experimental, *cuasi* experimental o no experimental. De acuerdo con los objetivos de este trabajo, sólo dos diseños de estudio que validan un supuesto mediante la intervención de un investigador en las variables del experimento son relevantes:

- *Diseño experimental* es un diseño que se aplica cuando el investigador puede controlar todas las variables involucradas que otorgan validez a la hipótesis de la investigación y que se puede realizar una asignación aleatoria de la muestra (Gribbons & Herman, 1996).
- *Diseño cuasi experimental* es un diseño que se aplica en ambientes semicontrolados, es decir, cuando el investigador no puede controlar todas las variables involucradas que otorgan validez a la hipótesis de la investigación o cuando no se puede realizar una asignación aleatoria de la muestra (Gribbons & Herman, 1996), por lo tanto, la validez de los resultados se ve afectada por variables internas o externas al experimento, esto se traduce en problemas de interpretación.

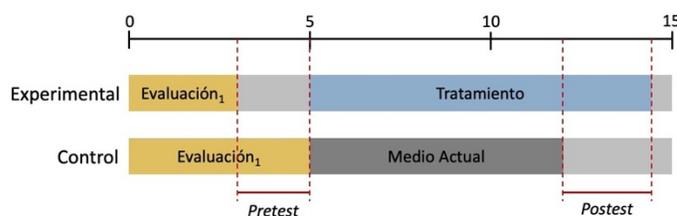
Un aspecto importante de la investigación experimental es el *aspecto temporal* del estudio, este aspecto se divide en dos posibilidades:

- *Estudio transversal* es un tipo de estudio observacional en el que se recolecta un conjunto de datos, con el objetivo de estudiar a una población en un determinado punto en el tiempo. Este estudio permite observar las relaciones que existen entre muchas variables de interés sin alterar el entorno natural donde se produce el fenómeno; en otras palabras, se pueden modificar ciertas variables y tomar nuevos datos para contrastar los resultados, incluso si los estudios se realizaron en distintos periodos de tiempo. Sin embargo, este tipo de estudio no es capaz de proporcionar información sobre la relación de causalidad de las variables, carece de validez interna (Rice, 1997).
- *Estudio longitudinal* es un tipo de estudio observacional en el cual se recolecta un conjunto de datos de la misma muestra y de forma repetitiva durante un periodo de tiempo prolongado (Rice, 1997). Este estudio proporciona datos sobre los cambios de la muestra y su tipo de cambio durante el experimento; además, permite a los investigadores determinar la secuencia de los eventos ocurridos, por consiguiente, puede proporcionar información más clara sobre la relación de causalidad de las variables (SET Sociedad Española de Toxicomanías, 2006).

Otro aspecto importante de acuerdo con el objetivo de este trabajo es el diseño del grupo de control, el *diseño pretest-postest con grupo de control no equivalente* es un tipo de diseño *cuasi* experimental que trabaja con muestras existentes de la población que se desean estudiar, es decir, no hay asignación aleatoria. Sin embargo, el investigador se debe dar a la tarea de asignar aleatoriamente la muestra disponible en grupos de investigación, con la finalidad de hacer los grupos de investigación equivalentes y evitar *variables extrañas* (las variables extrañas son variables que no se tomaron en cuenta en el momento del diseño del experimento y que influyen de alguna manera en la interacción entre la variable independiente y la variable dependiente, en otras palabras, influye en el resultado del experimento).

De acuerdo con la literatura, el diseño *pretest-postest* con grupo de control no equivalente indica que el tratamiento aplicado al grupo experimental tuvo efecto si y sólo si la diferencia entre los grupos experimental y control (*postest*) es mayor a la diferencia que existía inicialmente entre estos (*pretest*) (Huertas & Pantoja, 2016). En otras palabras, la medida *pretest* indica la diferencia de conocimiento que existía entre los grupos antes de introducir el tratamiento y la medida *postest* indica la diferencia que existe entre los grupos después de

introducir el tratamiento, por lo tanto, el tratamiento tuvo efecto si *posttest* es mayor a *pretest*. La Figura 31 expone la manera de evaluar la medida *pretest-posttest* de acuerdo con la literatura.



*Figura 31 Evaluación pretest-posttest*

El grupo de control y experimental quedan descritos de la siguiente manera:

- *Grupo experimental* que se somete al tratamiento del experimento con el objetivo de observar la manera en que se comportan las variables dependientes (VD) con los cambios hechos en las variables independientes (VI). En general, este grupo se acompaña de un grupo de control para aumentar la validez interna o la causalidad del efecto del tratamiento.
- *Grupo de control* que no se somete al tratamiento del experimento, sino que tiene la finalidad de servir de punto de contraste para el grupo experimental, la diferencia de los resultados entre ambos grupos es atribuida a los efectos de tratamiento aplicado.

Con base en lo anterior, se ha seleccionado un *estudio cuasi experimental longitudinal con diseño pretest-posttest con grupo de control no equivalente*, esto se debe a lo siguiente:

- En la literatura, se ha encontrado que predominan los estudios *cuasi* experimentales para evaluar sistemas educativos incluyendo los sistemas *CALL* (Ferreira, Vine, & Elejalde, 2015), esto se debe a la naturaleza de las muestras preexistentes y al control limitado sobre todas las variables que influyen en el aprendizaje de las personas.
- El tiempo es otra variable de suma importancia para tener en cuenta en la evaluación de un sistema de aprendizaje, además, el método del presente trabajo tiene como fundamento la *teoría de la curva del olvido* donde la retención de la información (aprendizaje) está determinada por el número de repasos y el tiempo entre estos repasos; con base en estas observaciones, se determina aplicar un estudio que permita dar seguimiento al aprendizaje de los estudiantes a través del tiempo. Este tipo de seguimiento sólo se puede realizar por medio de un estudio longitudinal.

- Además, se necesita comprobar que el efecto obtenido es originado por el tratamiento aplicado (causalidad), no obstante, un estudio *cuasi* experimental debilita esta característica del estudio. Con la finalidad de comprobar la causalidad y fortalecer el resultado obtenido del estudio *cuasi* experimental, se ha seleccionado un diseño *pretest-postest* con grupo de control no equivalente que permite dar corroborar la causalidad del tratamiento mediante la medición del estado inicial de los grupos contra su estado final.
- Por último, un estudio necesita una validación estadística que permita conocer si las conclusiones obtenidas son estadísticamente significativas, esto se logra mediante una *prueba de hipótesis de la diferencia de medias* de acuerdo con los objetivos del presente trabajo.

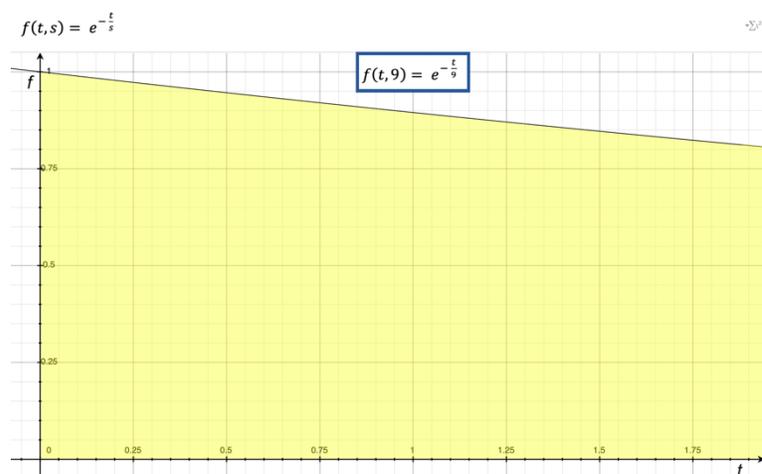
#### ***4.3.3 Estudio cuasi experimental longitudinal con diseño pretest-postest con grupo de control no equivalente***

*Objetivo:* determinar el efecto que tiene el uso del sistema SAIC desarrollado mediante el *método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera* en el aprendizaje de los estudiantes de una lengua extranjera.

*Hipótesis:* el uso del sistema SAIC con base en el *método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera* supone un mayor aprendizaje de vocabulario de una lengua extranjera en comparación con un sistema de cómputo de aprendizaje de vocabulario por traducción.

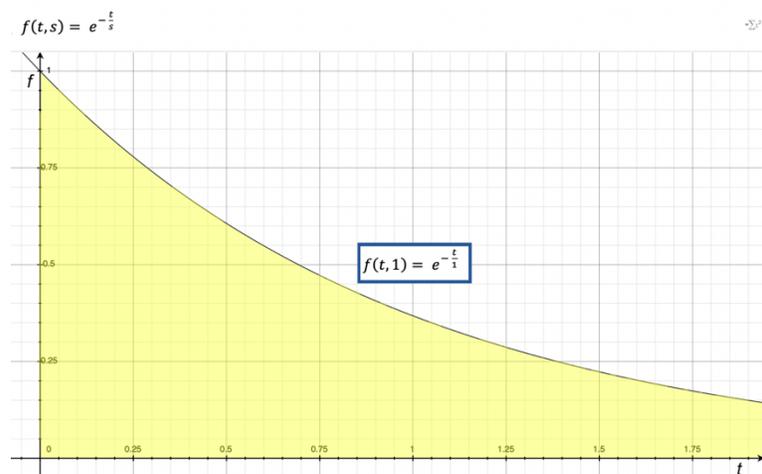
Las *variables* se dividen en dos: variables independientes (VI), corresponde a la interacción del estudiante con el sistema SAIC; y variable dependiente (VD), representa al aprendizaje de vocabulario del estudiante. Las *variables independientes* (VI) están compuestas por las siguientes actividades que realiza el estudiante dentro del sistema SAIC: lectura de textos, ampliación de vocabulario y repaso de vocabulario. La *variable dependiente* (VD) es la cantidad de palabras desconocidas que el estudiante ha logrado aprender durante las actividades que realizó en el sistema SAIC. Este aprendizaje se determina mediante la teoría de la curva del olvido con las variables: retención, tiempo y repetición. También se debe monitorear si las repeticiones espaciadas se realizan en tiempo como lo recomienda el algoritmo de repetición espaciada incorporado en el sistema SAIC.

De acuerdo con la teoría de la curva del olvido, un mayor número de repeticiones espaciadas correctas y consecutivas ( $s$ ) significan que la información se ha retenido en la memoria con más fuerza ( $f$ ) y durante más tiempo ( $t$ ) (Figura 32) en comparación con un menor número de repeticiones espaciadas correctas y consecutivas (Figura 33).



**Figura 32** Gráfica de la ecuación de la curva del olvido con nueve repeticiones espaciadas

**Nota.** Se puede observar un decremento suave de la fuerza de retención ( $f$ ) a través del tiempo ( $t$ ) por el efecto de nueve repeticiones espaciadas correctas y consecutivas ( $s=9$ ).



**Figura 33** Gráfica de la ecuación de la curva del olvido con una repetición espaciada

**Nota.** Se puede observar un decremento pronunciado de la fuerza de retención ( $f$ ) a través del tiempo ( $t$ ) por el efecto de una repetición espaciada correcta y consecutiva ( $s=1$ ).

En la literatura, se observa que un estudio sobre el aprendizaje humano requiere un periodo mínimo de un mes, se puede extender este periodo de acuerdo con el tiempo disponible y los objetivos del estudio. En este estudio se dispone de un periodo de un mes (mes hipotético de 32 días), además, se contempla que una buena participación de los estudiantes permite establecer un *indicador de aprendizaje* en 9 repeticiones espaciadas ( $s=9$ ) para clasificar una *palabra desconocida* a *palabra aprendida* (ver Tabla 14). Sin embargo, se puede utilizar una variante: si la participación de los estudiantes resulta escasa, entonces se establece el *indicador de aprendizaje* en el número máximo de repeticiones espaciadas que ha logrado el estudiante para cada palabra desconocida ( $s<9$ ) para clasificarla de *palabra desconocida* a *palabra en proceso de aprendizaje*.

*Tabla 14* Gráfica que determina el indicador de aprendizaje de acuerdo con el periodo disponible para la evaluación

Número de repasos ( $s$ )	Retención de la memoria ( $p$ )	$-s$	$\ln p$	Intervalo de tiempo ( $t$ ) (días)	Acumulado de $t$ (días)
1	0.5	-1	-0.6931	0.6931	0.6931
2	0.5	-2	-0.6931	1.3862	2.0793
3	0.5	-3	-0.6931	2.0794	4.1587
4	0.5	-4	-0.6931	2.7725	6.9312
5	0.5	-5	-0.6931	3.4657	10.3969
6	0.5	-6	-0.6931	4.1588	14.5557
7	0.5	-7	-0.6931	4.8520	19.4077
8	0.5	-8	-0.6931	5.5451	24.9528
9	0.5	-9	-0.6931	6.2383	31.1911

**Nota.** Una vez que se establece el periodo del estudio, se determina el *indicador de aprendizaje* para clasificar las palabras desconocidas que repasa el estudiante en: *palabras aprendidas*, se toma el número máximo de repeticiones espaciadas correctas y consecutivas que se pueden realizar durante el periodo de estudio (como lo muestra esta tabla,  $s=9$  proporciona una cobertura de 31.1911 días de estudio); o *palabras en proceso de aprendizaje*, tomando el número máximo de repeticiones espaciadas que logra el estudiante en cada palabra desconocida.

El *instrumento de recolección* que se utiliza para dar seguimiento al progreso de aprendizaje de los estudiantes es el sistema SAIC. Este sistema tiene la capacidad de recolectar toda la información que los estudiantes vayan generando a partir de sus sesiones de estudio. Sin embargo, se presentan diferencias en la experiencia de los estudiantes que utilizan este sistema debido a la naturaleza del experimento. Por un lado, el grupo experimental tiene a su disposición todas las funcionalidades del sistema SAIC descritas en el método del presente trabajo. Por otro lado, el grupo de control tiene a su disposición una versión del sistema SAIC con funciones reducidas, esta versión desactiva funciones como los algoritmos de recomendación de textos y

repetición espaciada, así como la implementación de signo lingüístico en las tarjetas de memoria; con la finalidad que la funcionalidad de esta versión reducida se parezca más a las funciones proporcionadas por el sistema competidor más cercano *ReadLang* desarrollado y distribuido por Steve Ridout (2021).

En 2019, se realizó una *prueba piloto* sobre la ecuación de la curva del olvido que se utiliza en el presente trabajo, propuesta por Hu y Ogihara (2011); con la finalidad de observar las variables que intervienen en el aprendizaje de un conjunto de palabras en inglés por parte de dos grupos de estudio (control y experimental) y determinar si la ecuación seleccionada es idónea o requiere un ajuste. Esta prueba se documentó en el artículo *un experimento para determinar el efecto de la curva del olvido en el aprendizaje del signo lingüístico* (Peralta-Quintero, Castro-Sánchez, González-Serna, & Nieto-Benítez, 2019) que forma parte de las actividades académicas adicionales de este trabajo de tesis (ver Figura 34), se concluyó que la ecuación de la curva del olvido elegida se aproxima a los resultados de aprendizaje del conjunto de palabras en inglés obtenidos en la evaluación realizada en el artículo, también se determinó que un estudio para evaluar un método de aprendizaje requiere de recolectar varias muestra de datos durante un periodo mayor que las dos semanas empleadas en el experimento del artículo. Además, se encontró que los grupos de control y experimental necesitan una muestra más grande de estudiantes que los seis estudiantes que participaron en el artículo, por último, se observó que la participación o tiempo de interacción de los estudiantes con el método de aprendizaje es una variable relevante para tener en cuenta. Las conclusiones obtenidas en dicho artículo fueron tomadas en cuenta para construir este estudio para el sistema SAIC.

## Un experimento para determinar el efecto de la curva del olvido en el aprendizaje del signo lingüístico

Carlos-Alberto Peralta-Quintero\*, Noé-Alejandro Castro-Sánchez\*  
Juan-Gabriel González-Serna\* Kenia Nieto-Benítez\*

\*Tecnológico Nacional de México / CENIDET, Cuernavaca, MOR 62490 México  
(e-mail: cqintero.xzv@icloud.com; {ncastro, gabriel, kenianieto}@cenidet.edu.mx)

**Resumen** La teoría de la curva del olvido de Ebbinghaus está fundamentada en el aprendizaje utilizando la memoria pura, sin asociaciones mentales. Este artículo se centra en un experimento para observar los efectos de la teoría de Ebbinghaus orientada hacia un enfoque asociativo, en particular, hacia el aprendizaje del signo lingüístico. El experimento consistió en dos grupos que realizaron repases espaciados para determinar el impacto que tiene el contenido y cantidad de material de estudio en el aprendizaje, aplicando la teoría de Ebbinghaus. Los resultados obtenidos no fueron suficientes para determinar la influencia del contenido del material de estudio en el aprendizaje. Por otro lado, se determinó que a mayor cantidad de material de estudio se necesita más tiempo para aprenderlo. Además, la teoría de Ebbinghaus ha funcionado correctamente en un enfoque asociativo.

**Palabras clave:** Curva del olvido, Ebbinghaus, Memoria, Signo lingüístico, Tarjetas de memoria, Repetición espaciada.

### 1. INTRODUCCIÓN

La memoria a corto plazo tiene una capacidad de almacenamiento y duración limitados, por el contrario, la memoria a largo plazo posee un almacenamiento relativamente permanente y su capacidad es vasta (Schimanke, Mertens, Ribbers y Vornberger, 2015).

En relación con lo anterior, el psicólogo alemán Hermann Ebbinghaus (1885) realizó un experimento con sílabas fáciles de recordar, pero sin significado. Este experimento tuvo como objetivo eliminar las asociaciones mentales que las palabras, frases u oraciones pudieran causar en la memoria, es decir, empleó un aprendizaje utilizando la memoria pura. No obstante, fuera de un ambiente controlado, el aprendizaje se lleva a cabo mediante asociaciones mentales; por ejemplo, cuando una persona aprende su lengua materna. Como resultado, Ebbinghaus publicó un mapeo sobre el decremento a través del tiempo de la capacidad de retención de la información de reciente adquisición, actualmente conocido como la teoría de la *curva del olvido* (Chmiel y Schubert, 2018).

Con respecto al aprendizaje de la lengua, Saussure, Bally y Alonso (2008) mencionan que la lengua existe en la colectividad de las personas, como si de un diccionario se tratara. Por tanto, si una persona quiere aprender la lengua de esa colectividad, necesitaría aprender dicho diccionario. Sin embargo, este enfoque es simplista, el mismo Saussure *et al.*

(2008) explican que no corresponde un nombre y una cosa, sino un concepto y una imagen acústica: el *signo lingüístico*.

Si bien, en la literatura, se pueden encontrar propuestas para aplicar o ajustar la teoría de la curva del olvido hacia un aprendizaje asociativo, no se ha obtenido un resultado concluyente que proporcione una ecuación ajustada hacia dicho tipo de aprendizaje.

Este artículo se centra en un experimento para observar los efectos de la teoría de la curva del olvido con variaciones en el contenido y cantidad de material de estudio, con la finalidad de enfocar la teoría de Ebbinghaus hacia un aprendizaje asociativo, en particular, hacia el aprendizaje del signo lingüístico. Se utiliza (1), propuesta por Hu y Ogihara (2011), para implementar la teoría de la curva del olvido en el presente experimento.

### 2. TRABAJOS RELACIONADOS

En la literatura, la *repetición espaciada* se define como el estudio de la información en intervalos de tiempo establecidos, estos intervalos pueden ser calculados mediante la teoría de la curva del olvido de Ebbinghaus. Por ejemplo, Zheng (2015) propuso un sistema de aprendizaje de vocabulario, el cual supone que el usuario estudie una lista de palabras mediante repeticiones espaciadas. Hsiao, Cheung, Jiang y Yu (2016) propusieron un sistema que programa sesiones de estudio espaciadas para memorizar el vocabulario del examen de inglés IELTS. Swart (2018) realizó un estudio longitudinal donde

Figura 34 Primera página del artículo “Un experimento para determinar el efecto de la curva del olvido en el aprendizaje del signo lingüístico” (Peralta-Quintero, Castro-Sánchez, González-Serna, & Nieto-Benítez, 2019)

## Capítulo 5 Evaluación

Se realizó un experimento que consistió en la participación de dos grupos de estudiantes en las actividades de lectura de textos, ampliación de vocabulario y repaso de vocabulario dentro del sistema SAIC. Estas actividades fueron monitoreadas continuamente y muestreadas de forma semanal durante el periodo de evaluación, con la finalidad de obtener información acerca del aprovechamiento en el aprendizaje de los grupos de estudio, contrastar la información y proporcionar un resultado sobre el beneficio de usar el sistema SAIC implementado con base en el *método para la adquisición y reforzamiento de vocabulario para el aprendizaje de una lengua extranjera*.

Cabe señalar que los estudiantes del grupo experimental tuvieron acceso total a las funciones del sistema SAIC para las actividades del experimento, como lo describe el método de este trabajo; mientras que los estudiantes del grupo de control tuvieron acceso limitado al sistema, con el objetivo que el sistema se pareciera al servicio ofrecido por el sitio *readlang.com* (Ridout, 2021) para poder contrastar los beneficios que tiene el nuevo sistema frente a su competidor más cercano.

En relación con el sitio *readlang.com* (Ridout, 2021), se tomó como punto de referencia debido a su semejanza con el sistema SAIC, principalmente porque ambos sistemas están orientados hacia el aprendizaje de vocabulario con base en la lectura de textos y uso de tarjetas de memoria como material de estudio. Por lo tanto, una comparación entre las funciones y los beneficios de estos sistemas proporcionó una diferencia en el aprendizaje de los estudiantes que se puede atribuir a las características adicionales que propone el método del presente trabajo.

Por otro lado, se utilizó un estudio *pretest-postest* con grupo de control no equivalente para validar la propiedad de causalidad del experimento, este estudio tiene un diseño *cuasi* experimental que trabaja con muestras existentes de la población que se desea estudiar, es decir, no hay asignación aleatoria. Sin embargo, se realizó una asignación aleatoria simple al grupo de control y experimental mediante el registro abierto y un algoritmo de aleatoriedad. Por último, se empleó una prueba de hipótesis de la diferencia de medias para conocer si los resultados obtenidos en el estudio son estadísticamente significativos.

## 5.1 Corpus de textos en inglés

El corpus de textos del sistema SAIC en producción estuvo conformado por 746 textos cortos en inglés de distintas fuentes de información. A continuación, la Tabla 15 proporciona detalles sobre las fuentes proveedoras de los textos.

*Tabla 15 Fuente de los textos del corpus del sistema SAIC*

<b>Fuente</b>	<b>Autor</b>	<b>Tipo</b>	<b>Cantidad</b>
MCTest Project	Richardson, Burges y Renshaw (2013)	Historias cortas	660
The Little Prince	De Saint-Exupery (2001)	Fragmentos del cuento	14
The Canterville Ghost and Other Stories	Wilde (2001)	Fragmentos de cuento	20
Simple English Wikipedia	Wikipedia Foundation (2021)	Fragmentos de artículos	35
Test English	Test-English.com (2021)	Anécdotas	12
British Council - Learn English	British Council (2021)	Anécdotas	5
<b>Total</b>			<b>746</b>

Estos textos fueron clasificados por su nivel de complejidad de acuerdo con el MCER (Marco común europeo de referencia para las lenguas) por medio de la herramienta *TextAnalyzer* (RoadToGrammar, 2019), con la finalidad de proporcionar un conjunto de textos a los estudiantes de acuerdo con su dominio en el idioma inglés. La Tabla 16 muestra la cantidad de textos disponibles de acuerdo con el nivel de complejidad propuesto por el MCER.

*Tabla 16 Cantidad de textos del sistema SAIC por complejidad de acuerdo con el MCER*

<b>Nivel de complejidad</b>	<b>Cantidad</b>
A1	53
A2	90
B1	374
B2	166
C1	53
C2	10
<b>Total</b>	<b>746</b>

Después de la clasificación de textos por su nivel de complejidad, se realizó el preprocesamiento (normalización, lematización y tokenización) descrito en el *algoritmo de recomendación de textos* para obtener los componentes de las lecturas, por último, se insertaron en la base de datos del sistema SAIC las lecturas originales, así como los componentes obtenidos en el preprocesamiento realizado.

## 5.2 Participantes

La evaluación tuvo una participación de 115 estudiantes de nivel medio superior de una misma institución educativa (Colegio de Bachilleres), sin embargo, sólo 66 estudiantes completaron las cuatro semanas de evaluación, comprendidas del 12 de abril al 9 de mayo del 2021. Por lo tanto, se realizó el análisis de las actividades con los 66 estudiantes divididos en grupo de control y experimental, con 33 estudiantes cada uno. Derivado de la pandemia por covid19, todos los participantes en este estudio fueron contactados vía remota y las actividades fueron realizadas mediante el servicio Web del sistema SAIC. A continuación, se muestra información relevante sobre los estudiantes participantes.

- *Género*: se tuvo una participación de 62.12% de mujeres y 37.88% de hombres como se muestra en la Figura 35.



*Figura 35 Participación de estudiantes por género*

- *Edad*: la mayoría de los estudiantes participantes reportó entre 16 y 17 años con un total del 79.04% de la participación como lo muestra la Figura 36.



Figura 36 Participación de estudiantes por edad

- *Estilo de aprendizaje dominante/secundario*: se aplicó el cuestionario de estilos de aprendizaje del modelo VAK (Secretaría de Educación Pública, 2004) para identificar el estilo de aprendizaje de los estudiantes. El estilo más común fue el flexible *auditivo, kinestésico y visual* con 11 estudiantes como lo muestra la Figura 37. Sin embargo, se realizó un acumulado por estilo dominante de todos los estudiantes para determinar qué material didáctico es mejor para trabajar con estos grupos y se observó que los tres estilos están casi equilibrados como se muestra en la Figura 38. Por lo tanto, se recomienda utilizar material didáctico diverso para darle cobertura a los tres canales de aprendizaje.

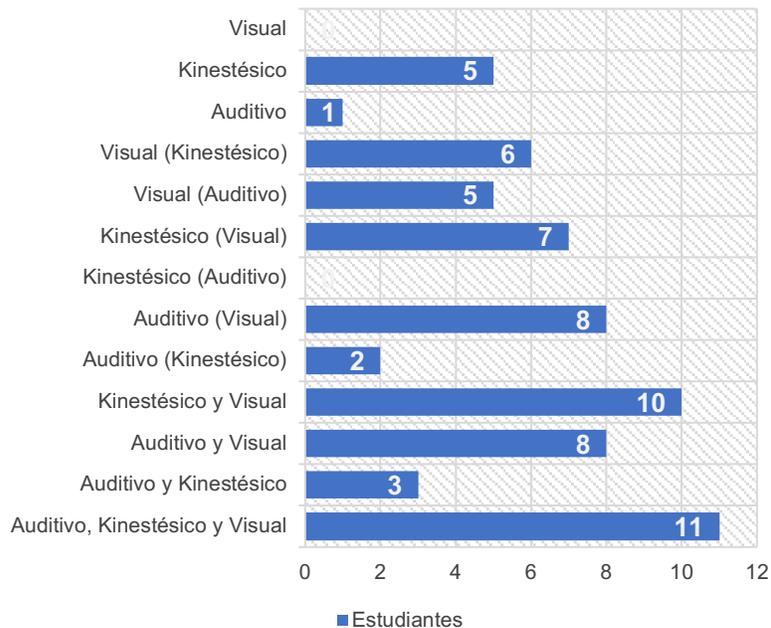
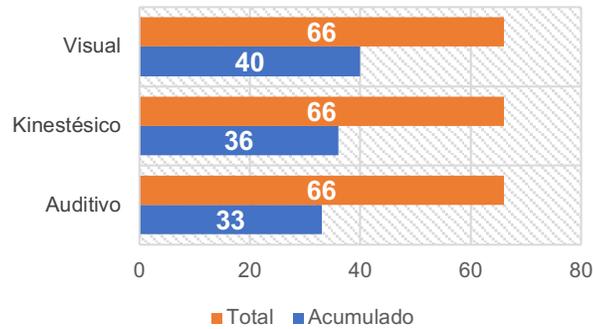
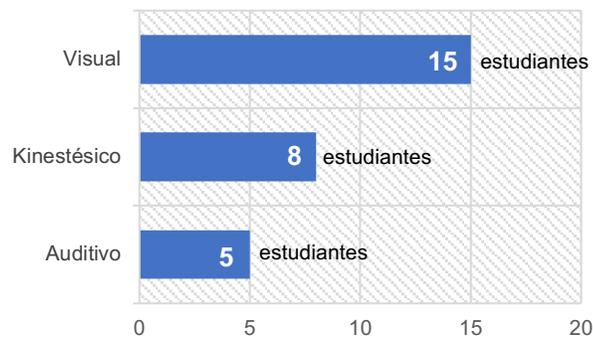


Figura 37 Participación de estudiantes por estilo de aprendizaje dominante/secundario



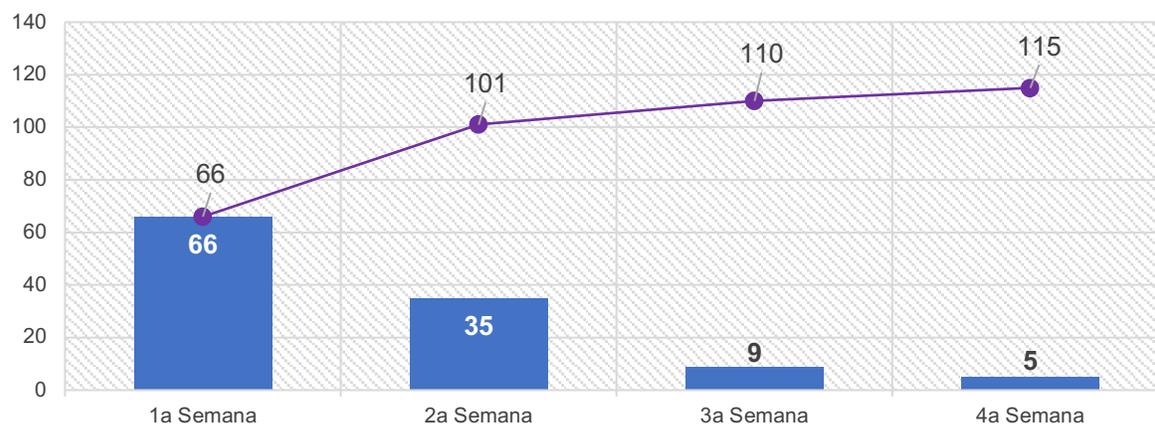
*Figura 38 Participación de estudiantes por estilo de aprendizaje dominante acumulado*

- *Estilo de aprendizaje secundario*: se observó un estilo de aprendizaje dentro de paréntesis en la Figura 37, este hace referencia al estilo de aprendizaje secundario. En la Figura 39, se observa que la mayoría de los estudiantes con un estilo de aprendizaje secundario presenta la facilidad de desarrollar su canal *visual*.



*Figura 39 Participación de estudiantes por estilo de aprendizaje secundario*

- *Nivel de dominio del idioma inglés*: los estudiantes participantes contaban con un nivel de dominio A1 de acuerdo con la información compartida por su docente a cargo.
- *Registro*: los estudiantes se pudieron registrar durante las cuatro semanas de evaluación como se presenta en la Figura 40. Sin embargo, el análisis del estudio sólo se realizó tomando en cuenta los 66 estudiantes que se registraron durante la 1ª semana de evaluación.



*Figura 40 Participación de estudiantes por registro en el sistema SAIC durante el periodo de evaluación*

Esta información estadística acerca de los estudiantes se obtuvo de las muestras de datos semanales tomadas del sistema SAIC.

### **5.3 Presentación de introducción**

El día 25 de marzo del 2021 se realizó una presentación de introducción a los estudiantes participantes, con la finalidad de motivarlos en el aprendizaje del idioma inglés como segunda lengua mediante casos de oportunidad en el ámbito social, académico, laboral, entretenimiento, ciencia y tecnología. También se les proporcionó información del método del presente trabajo y sobre el sistema SAIC.

La Figura 41 muestra la videollamada donde se realizó la presentación de introducción, en la cual estuvieron presentes los estudiantes participantes, la docente a cargo, el director de tesis y el tesista investigador. Además, la Figura 42 muestra la portada de la presentación expuesta en dicha videollamada.

Por otro lado, se acordó realizar una documentación de las muestras semanales tomadas del sistema SAIC y presentarlas en dos reportes que dan seguimiento a los estudiantes en sus estadísticas generales, registro y actividades de aprendizaje, con el objetivo de entregar un reporte a su docente a cargo para hacer de su conocimiento el progreso de aprendizaje complementario a su clase de inglés y un segundo reporte para los fines académicos de la investigación del presente trabajo.

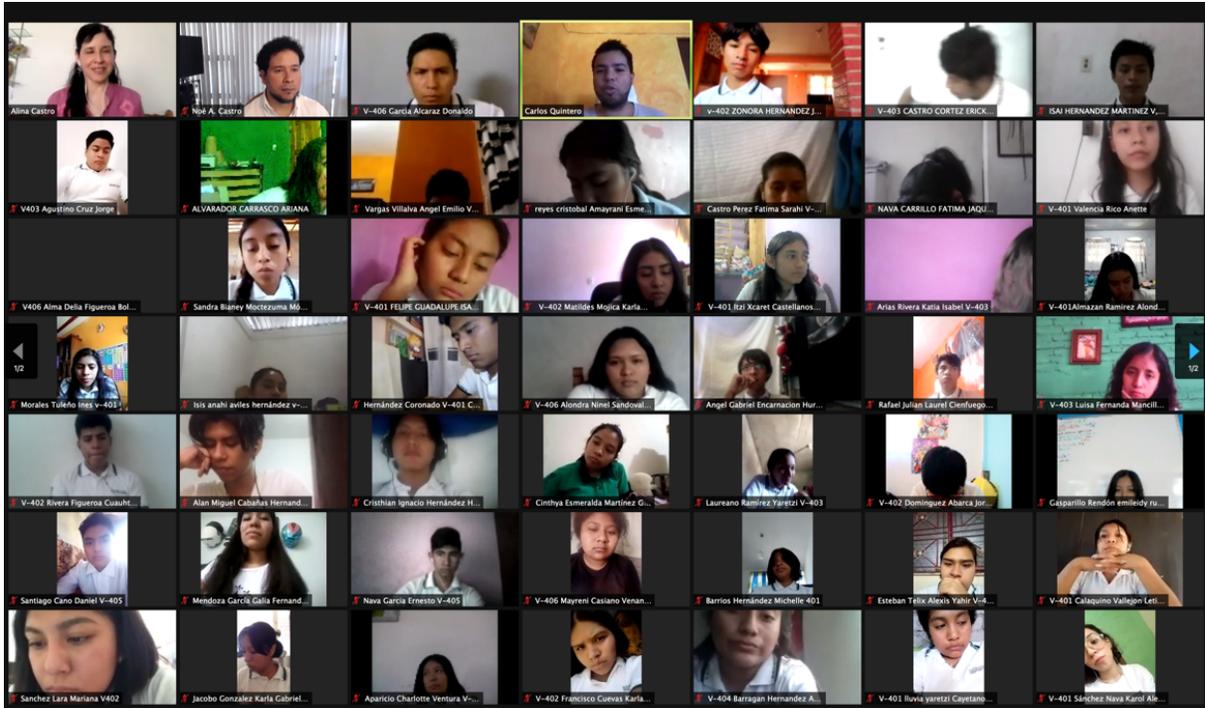


Figura 41 Participantes en la videollamada de la presentación de introducción

Figura 42 Portada de la presentación de introducción durante la videollamada

## 5.4 Competidor más cercano del sistema SAIC

De acuerdo con la investigación sobre productos relacionados elaborada en el presente trabajo, el sitio *readlang.com* (Ridout, 2021) proporciona un servicio para el aprendizaje de una lengua extranjera similar al sistema SAIC, ambos sistemas tienen como base del aprendizaje la

lectura de textos y el uso de tarjetas de memoria. A continuación, se describen las principales funciones que provee el sitio *readlang.com* (Ridout, 2021) clasificadas por características contrastadas y no contrastadas contra el sistema SAIC.

- Características contrastadas
  - *Online eReader* es un lector de textos que proporciona traducción de palabras y oraciones para auxiliar al usuario en la lectura de textos en una lengua extranjera. Este lector guarda automáticamente las palabras u oraciones traducidas.
  - *Flashcards* es una actividad didáctica de tarjetas de memoria que proporciona aprendizaje tradicional, es decir, la memorización de una palabra en lengua extranjera con su traducción a la lengua materna del usuario. Es importante señalar que la última versión del sitio incorpora un algoritmo de repetición espaciada, sin embargo, no se proporciona información sobre dicho algoritmo ni se muestra la organización de palabras para su estudio.
- Características no contrastadas
  - *Web reader* es un complemento para el navegador Safari que permite agregar una palabra al vocabulario de *readlang.com* (Ridout, 2021) desde cualquier página Web. Actualmente el complemento está deshabilitado.
  - *Customizable dictionary* es una funcionalidad que permite buscar el significado de una palabra traducida mediante cualquier diccionario en línea que sea compatible.
  - *Vocab manager* es una característica que permite la administración de palabras y oraciones agregadas al vocabulario del usuario.
  - *Video player* es una característica que permite afinar el oído mediante videos y audios en lengua extranjera con subtítulos.

Se adaptaron las funcionalidades del sistema SAIC para que el grupo de control tuviera las características contrastadas del sitio *readlang.com* (Ridout, 2021), con el objetivo de comparar las funcionalidades adicionales propuestas en el presente trabajo y determinar el beneficio que estas le proporcionan a los estudiantes de una lengua extranjera.

## 5.5 Actividad de exploración

La actividad de exploración tiene dos objetivos: indagar en el conocimiento del estudiante a partir de una lista de palabras en lengua extranjera y amortiguar el problema de inicio en frío en el algoritmo de recomendación de textos implementado en el sistema SAIC, como se ha presentado en capítulos anteriores. La lista de diez palabras que se utilizó en esta actividad se obtuvo de manera aleatoria de los textos del sistema SAIC con nivel de complejidad igual que el nivel de dominio del estudiante, en este caso se utiliza la misma lista debido a que todos los estudiantes tienen un nivel de dominio A1.

En la actividad de exploración, el estudiante seleccionó las palabras que conocía de la lista de diez palabras presentada (ver Figura 43); las palabras seleccionadas se catalogaron como *palabras conocidas* y se utilizaron en el análisis del estudio como parte de la variable PC para calcular la medida *pretest-postest*, mientras que las palabras no seleccionadas se catalogaron como *palabras desconocidas* y se agregaron al vocabulario del estudiante de forma automática.

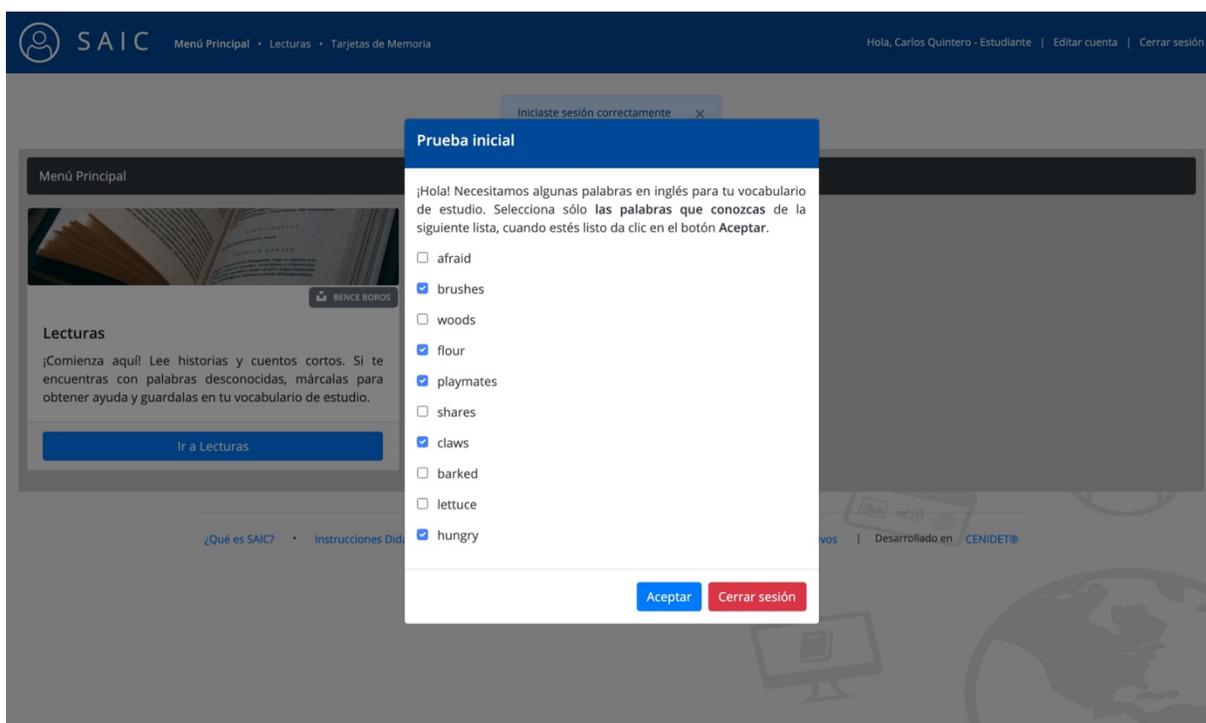


Figura 43 Selección de palabras conocidas en la actividad de exploración

## 5.6 Actividades de aprendizaje

Los estudiantes participantes realizaron las actividades: lectura de textos, ampliación de vocabulario y repaso de vocabulario dentro del sistema SAIC. Sin embargo, los estudiantes del

grupo de control tuvieron restricciones en las funciones proporcionadas por el sistema para estas actividades, con la finalidad que la experiencia de dicho grupo fuera lo más parecido a la funcionalidad ofrecida por el sitio *readlang.com* (Ridout, 2021). Las actividades de la evaluación se dividieron en cuatro semanas con las siguientes fechas programadas:

- 1ª semana del 12 al 18 de abril del 2021.
- 2ª semana del 19 al 25 de abril del 2021.
- 3ª semana del 26 de abril al 2 de mayo del 2021.
- 4ª semana del 3 al 9 de mayo del 2021.

En los siguientes apartados se describen las actividades de aprendizaje realizadas y las características que estuvieron disponibles para cada grupo de estudio.

### **5.6.1 Lectura de textos**

Los estudiantes tuvieron acceso a los textos del sistema SAIC que correspondieron con su nivel de dominio del idioma inglés, es decir, los estudiantes con nivel de dominio A1 tuvieron acceso a 53 textos cortos en inglés catalogados en dicho nivel. El sistema proporcionó las siguientes funciones a los estudiantes de acuerdo con su grupo de estudio:

- Lista de lecturas con filtros.
  - Grupo experimental: tuvo acceso a los filtros por *lecturas nuevas, lecturas leídas, lecturas recomendadas, lecturas subidas por el sistema y todas las lecturas*.
  - Grupo de control: tuvo acceso a los mismos filtros de lecturas que el grupo experimental, excepto al filtro de *lecturas recomendadas*, esto se debe a que el algoritmo de recomendación de textos estuvo desactivado para este grupo de estudio.
- Traducción de palabras desconocidas durante la lectura de un texto corto en inglés.
  - Grupo experimental y control: ambos grupos tuvieron esta funcionalidad.

La Figura 44 presenta la lista de lecturas con filtros del sistema SAIC para el grupo experimental mientras que la Figura 45 muestra la lista para el grupo de control, así como la Figura 46 expone el visor de lectura para ambos grupos de estudio.

SAIC Menú Principal · Lecturas · Tarjetas de Memoria · Grupos Hola, Carlos Quintero - Docente | Editar cuenta | Cerrar sesión

Menú Principal / Lecturas recomendadas

Todas Leídas + Nuevas + Recomendadas + Subidas Crear lectura

**Nuestras recomendaciones.**  
 Lista de lecturas recomendadas con base en tu vocabulario de estudio, una lectura marcada con 5 \* es altamente recomendada mientras que una lectura con 0 \* no es recomendada.

Once upon a time jimmy had a mother who told him that he was good at music.	3.3 *	A2
Paws the cat lives with the Jones family.	3.2 *	A2
In a far away land known as Board, there was a great man named Pawn.	2.4 *	A2
My name is Clyde.	2.4 *	A2
A guy with brown hair was named Josh.	2.3 *	A2
There once was a guy named Kevin and girl named Erin.	2.2 *	A2
One morning, Becca was putting on her shoes when she heard a knock at the door.	2.1 *	A2
Hi there!	2.8 *	A2
There were four men who all played basketball.	2.6 *	A2
A lamb named Lana woke up on an early Monday morning.	2.1 *	A2

< Prev 1 Next >

[¿Qué es SAIC?](#) · [Instrucciones Didácticas](#) · [Acerca de Nosotros](#) · [Contáctanos](#) · [FB: Sistemas Cognitivos](#) | Desarrollado en [CENIDET®](#)

SAIC Versión 0.8.1 (Compilación 20210424)

*Figura 44 Lista de lecturas con filtros para el grupo experimental*

SAIC Menú Principal · Lecturas · Tarjetas de Memoria Hola, Carlos Quintero - Estudiante | Editar cuenta | Cerrar sesión

Menú Principal / Lecturas

Todas Leídas + Nuevas + Subidas Crear lectura

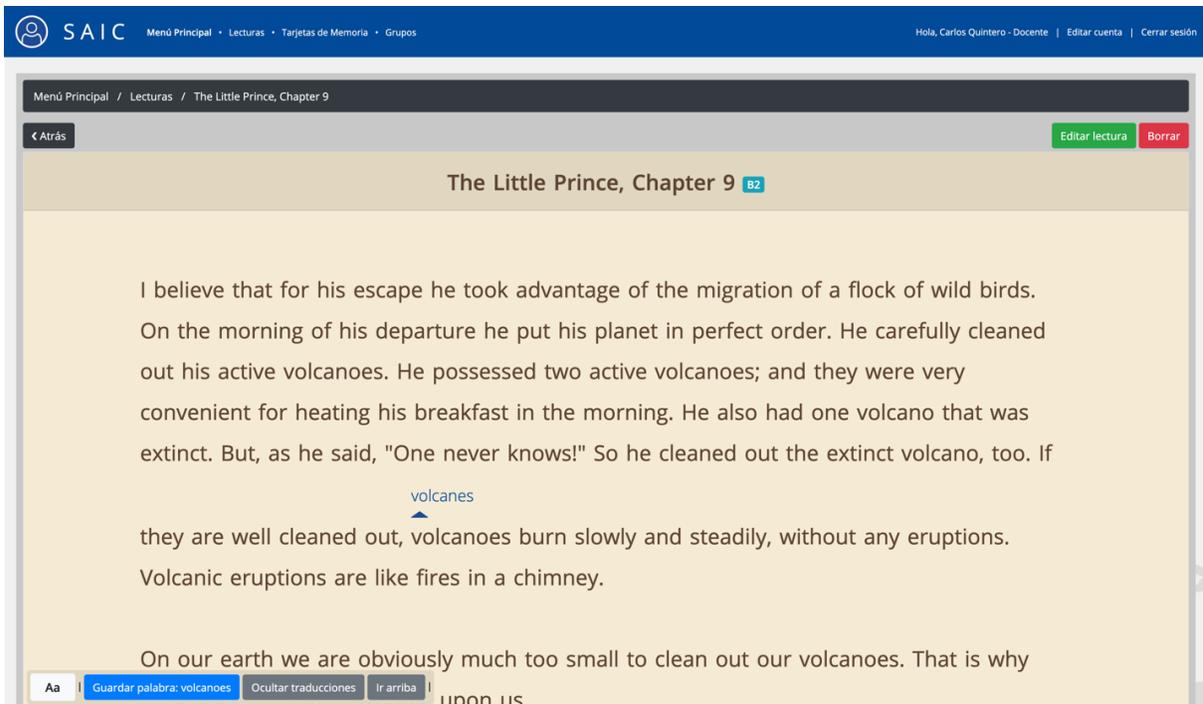
Pester came running into the room.	A2
There once was a guy named Kevin and girl named Erin.	A2
Nana had a bunch of money.	A2
I dont think I can win the race Tim told his parents over breakfast.	A2
Jake was walking to the park to play baseball with his friends.	A2
Jack and Mackenzie wanted to do something fun during their day off from school.	A2
Kelly woke up one morning and went outside to her back yard.	A2
One day, my family took a vacation to the beach.	A2
Timmy had the best grandmother ever.	A2
Luke was starting his first day of day care.	A2

< Prev 1 2 3 4 5 ... 9 Next >

[¿Qué es SAIC?](#) · [Instrucciones Didácticas](#) · [Acerca de Nosotros](#) · [Contáctanos](#) · [FB: Sistemas Cognitivos](#) | Desarrollado en [CENIDET®](#)

SAIC Versión 0.8.1 (Compilación 20210424)

*Figura 45 Lista de lecturas con filtros para el grupo de control*



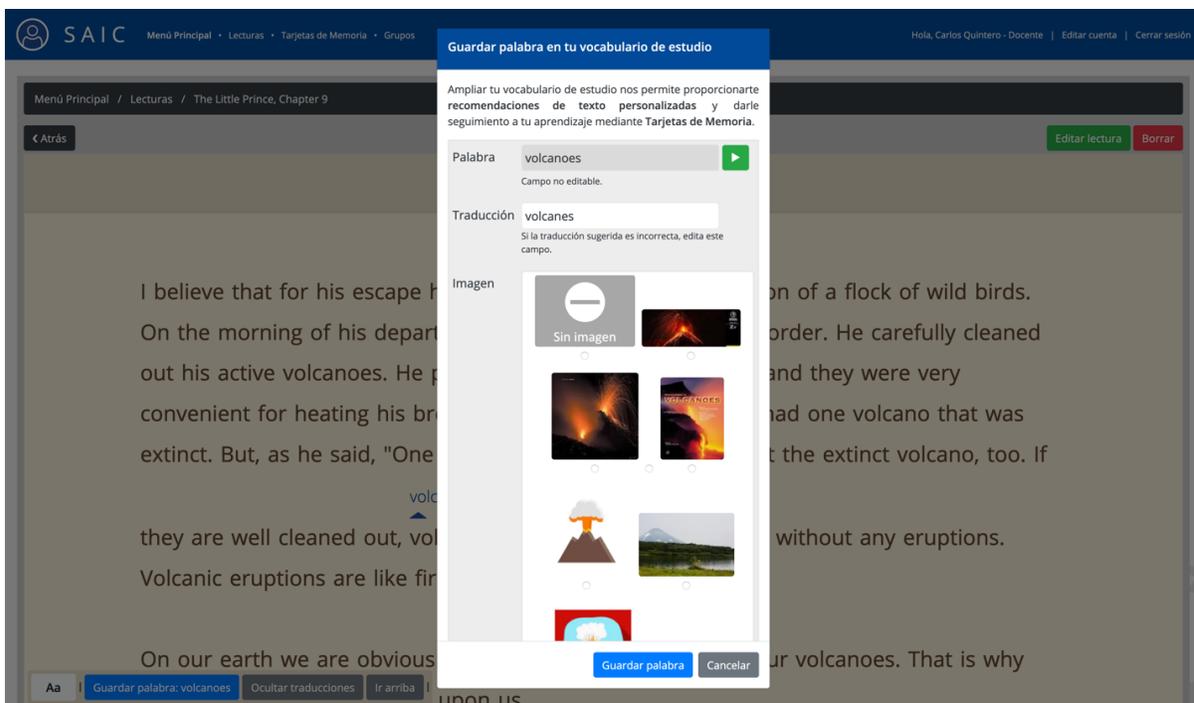
*Figura 46* Visor de lectura para ambos grupos de estudio

### **5.6.2 Ampliación de vocabulario**

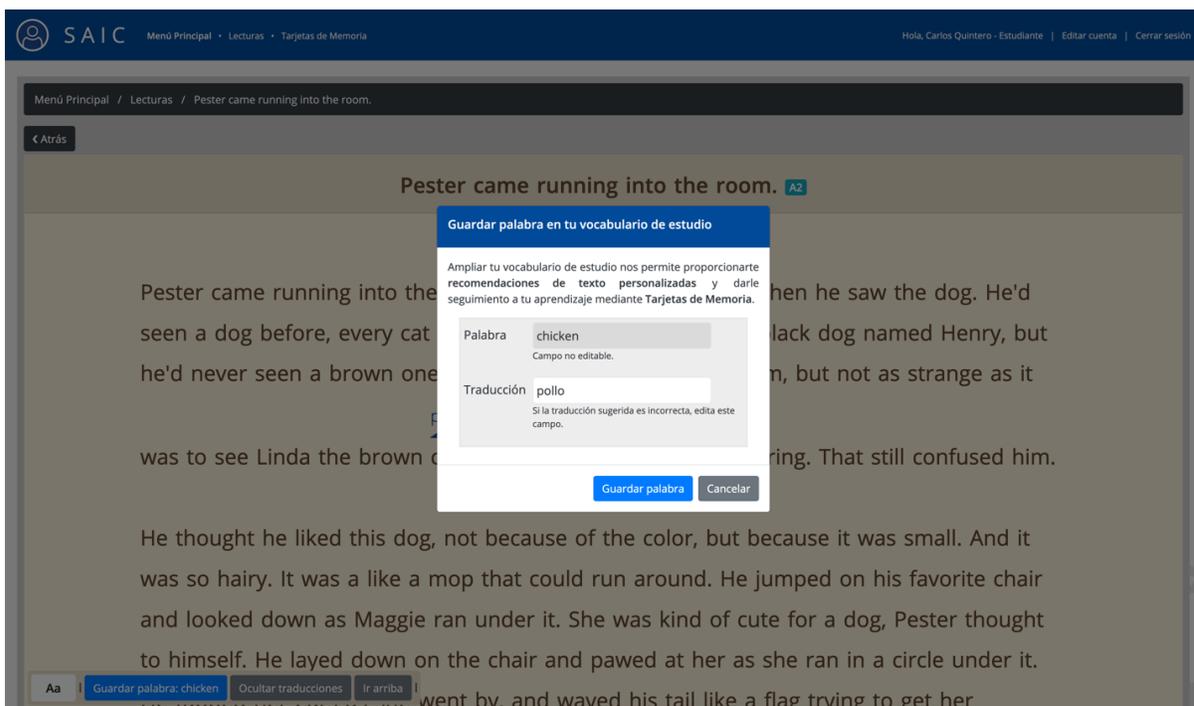
Los estudiantes tuvieron disponible la opción de guardar en su vocabulario de estudio las palabras desconocidas que encontraran durante su lectura por medio de las siguientes funciones:

- Marcado de palabras desconocidas durante la lectura.
  - Grupo experimental y de control: ambos grupos tuvieron esta funcionalidad.
- Pronunciación de la palabra desconocida.
  - Grupo experimental: tuvo acceso a un botón para reproducir la pronunciación de la palabra desconocida que se iba a guardar en su vocabulario.
  - Grupo de control: no tuvo acceso a esta funcionalidad.
- Asociación de una palabra desconocida con una imagen que, para el estudiante, mejor la representa.
  - Grupo experimental: tuvo acceso a una galería para seleccionar y adjuntar una imagen a la palabra desconocida que se iba a guardar en su vocabulario.
  - Grupo de control: no tuvo acceso a esta funcionalidad.

La Figura 47 muestra la vista para guardar una palabra desconocida en el vocabulario del estudiante perteneciente al grupo experimental mientras que la Figura 48 muestra la vista para un estudiante perteneciente al grupo de control.



*Figura 47 Vista para guardar una palabra en el vocabulario para el grupo experimental*



*Figura 48 Vista para guardar una palabra en el vocabulario para el grupo de control*

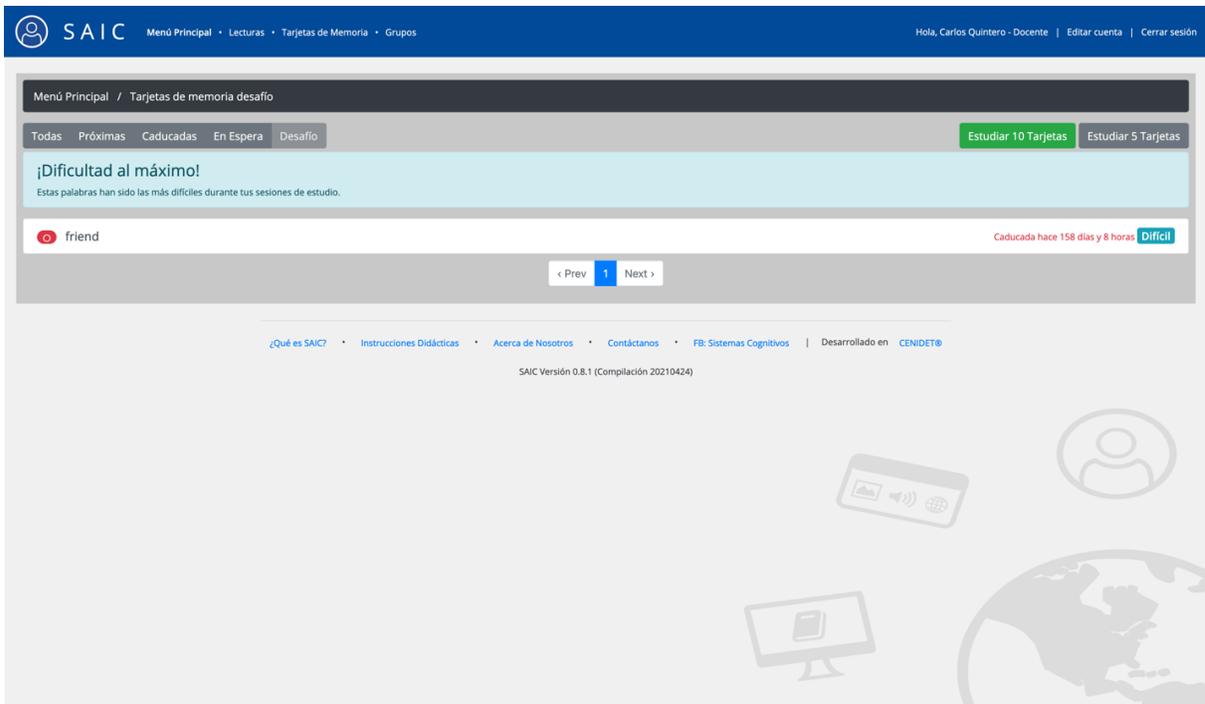
### 5.6.3 Repaso de vocabulario

Los estudiantes tuvieron acceso a una dinámica de tarjetas de memoria para estudiar las palabras desconocidas que agregaron a su vocabulario durante su lectura, con el objetivo de adquirir o reforzar dichas palabras. Los estudiantes contaron con las siguientes funciones:

- Organización de tarjetas de memoria mediante repetición espaciada.
- Grupo experimental: tuvo acceso al seguimiento del aprendizaje de palabras desconocidas para la dinámica de tarjetas de memoria por medio del algoritmo de repetición espaciada.
- Grupo de control: no tuvo acceso al seguimiento del aprendizaje de palabras desconocidas, el algoritmo de repetición espaciada estuvo desactivado para este grupo de estudio.
- Lista de tarjetas de memoria con filtros.
  - Grupo experimental: tuvo acceso a los filtros por *palabras próximas* (palabras para estudiar ese día), *palabras caducadas* (palabras que debieron estudiarse en días pasados), *palabras en espera* (palabras para estudiar en días siguientes), *palabras con desafío* (palabras particularmente difíciles de recordar para el estudiante) y *todas las palabras*.
  - Grupo de control: tuvo acceso sólo al filtro por *todas las palabras*, esto se debe a que el algoritmo de repetición espaciada estuvo desactivado para este grupo de estudio.
- Aprendizaje del signo lingüístico de una lengua extranjera mediante asociación visual, auditiva y escrita.
  - Grupo experimental: tuvo acceso al aprendizaje asociativo mediante una imagen relacionada y la pronunciación de la palabra en lengua extranjera, así como la validación escrita de la palabra en lengua extranjera que recordó el estudiante.
  - Grupo de control: tuvo acceso sólo al aprendizaje tradicional con tarjetas de memoria mediante la traducción de la palabra en lengua extranjera.
- Ayuda para recordar palabras en lengua extranjera durante los repastos con tarjetas de memoria.

- Grupo experimental: tuvo acceso a la sección de ayuda compuesta por una oración de contexto de la palabra desconocida (oración del texto donde se encontró la palabra por primera vez), traducción de la palabra desconocida a la lengua materna del estudiante y número de letras de la palabra desconocida.
- Grupo de control: no tuvo acceso a esta funcionalidad.

La Figura 49 presenta la lista de tarjetas de memoria para el grupo experimental mientras que la Figura 50 muestra la lista para el grupo de control. Además, la Figura 51 expone una tarjeta de memoria para el grupo experimentas mientras que la Figura 52 muestra una tarjeta para el grupo de control.



*Figura 49* Lista de tarjetas de memoria con filtros para el grupo experimental

SAIC Menú Principal · Lecturas · Tarjetas de Memoria Hola, Carlos Quintero - Estudiante | Editar cuenta | Cerrar sesión

Menú Principal / Tarjetas de memoria

Todas Estudiar 10 Tarjetas | Estudiar 5 Tarjetas

afraid	16 horas	Fácil
brushes	16 horas	Fácil
woods	16 horas	Fácil
flour	16 horas	Fácil
playmates	16 horas	Fácil
shares	16 horas	Fácil
claws	16 horas	Fácil
barked	16 horas	Fácil
lettuce	16 horas	Fácil
hungry	16 horas	Fácil

< Prev 1 Next >

[¿Qué es SAIC?](#) · [Instrucciones Didácticas](#) · [Acerca de Nosotros](#) · [Contáctanos](#) · [FB: Sistemas Cognitivos](#) | [Desarrollado en CENIDET®](#)  
 SAIC Versión 0.8.1 (Compilación 20210424)

Figura 50 Lista de tarjetas de memoria con filtro para el grupo de control

SAIC Menú Principal · Lecturas · Tarjetas de Memoria · Grupos Hola, Carlos Quintero - Docente | Editar cuenta | Cerrar sesión

Menú Principal / Tarjetas de memoria / Repaso de vocabulario

< Atrás

**Repaso de palabras en Inglés.** x

Observa la imagen y escucha la pronunciación de la palabra desconocida. Después, escribe la palabra que recuerdes en el campo de texto.  
 Nota: puedes apoyarte con la oración de contexto, la traducción al Español y el número de letras de la palabra.

Tarjeta 1 2 3 4 5

**¿Recuerdas la palabra?**

Observa



Escucha

▶

Pronunciar

Escribe

Palabra en Inglés

Siguiente Tarjeta

Ayuda

Ayuda 1: Oración de Contexto

He'd give the queen shorts as \_\_\_\_\_ presents.

Ayuda 2: Traducción Al Español

cumpleaños

Ayuda 3: Número de Letras

8 Letras

Figura 51 Tarjeta de memoria con aprendizaje asociativo para el grupo experimental

SAIC Menú Principal · Lecturas · Tarjetas de Memoria Hola, Carlos Quintero - Estudiante | [Editar cuenta](#) | [Cerrar sesión](#)

Menú Principal / Tarjetas de memoria / Repaso de vocabulario

[← Atrás](#)

**Repaso de palabras en Inglés.** ×  
Escribe la palabra que recuerdes en el campo de texto de acuerdo con la traducción mostrada.

Tarjeta 1 2 3 4 5

**¿Recuerdas la palabra?**

Traducción  
temeroso

Escribe

Palabra en Inglés

[Siguiente Tarjeta](#)

[¿Qué es SAIC?](#) · [Instrucciones Didácticas](#) · [Acerca de Nosotros](#) · [Contáctanos](#) · [FB: Sistemas Cognitivos](#) | Desarrollado en [CENIDET®](#)

SAIC Versión 0.8.1 (Compilación 20210424)



*Figura 52 Tarjeta de memoria con aprendizaje tradicional para el grupo de control*

# Capítulo 6 Resultados

Durante las cuatro semanas de la evaluación, se realizó una toma de muestra semanal de los datos generados por la interacción de los estudiantes con el sistema SAIC. Los datos muestreados fueron los siguientes:

- *Textos leídos* de la actividad lectura de textos.
- *Palabras agregadas al vocabulario* de la actividad ampliación de vocabulario.
- *Palabras estudiadas* de la actividad repaso de vocabulario.
- *Palabras en proceso de aprendizaje* de la actividad repaso de vocabulario.

La Tabla 17 muestra un resumen de los resultados de la participación e interacción de los estudiantes con el sistema SAIC. En esta tabla se puede observar que en todas las actividades hubo una mayor participación de los estudiantes del grupo de control en comparación con los estudiantes del grupo experimental, sin embargo, los estudiantes del grupo experimental que han participado han leído más textos, agregado más palabras a su vocabulario, estudiado más palabras y lograron más palabras en proceso de aprendizaje.

*Tabla 17 Resultados de la participación del grupo experimental y de control en las actividades de evaluación*

Datos muestreados	Grupo experimental		Grupo de control	
	<i>Participación*</i>	<i>Cantidad**</i>	<i>Participación*</i>	<i>Cantidad**</i>
Actividad de exploración	32	<b>3.697</b>	<b>33</b>	3.545
Textos leídos	10.75	<b>2.36</b>	<b>11.25</b>	2.27
Palabras agregadas al vocabulario	13	<b>8.78</b>	<b>15</b>	8.18
Palabras estudiadas	5.5	<b>7.59</b>	<b>7</b>	6.13
Palabras en proceso de aprendizaje	5	<b>6.64</b>	<b>6.5</b>	4.17

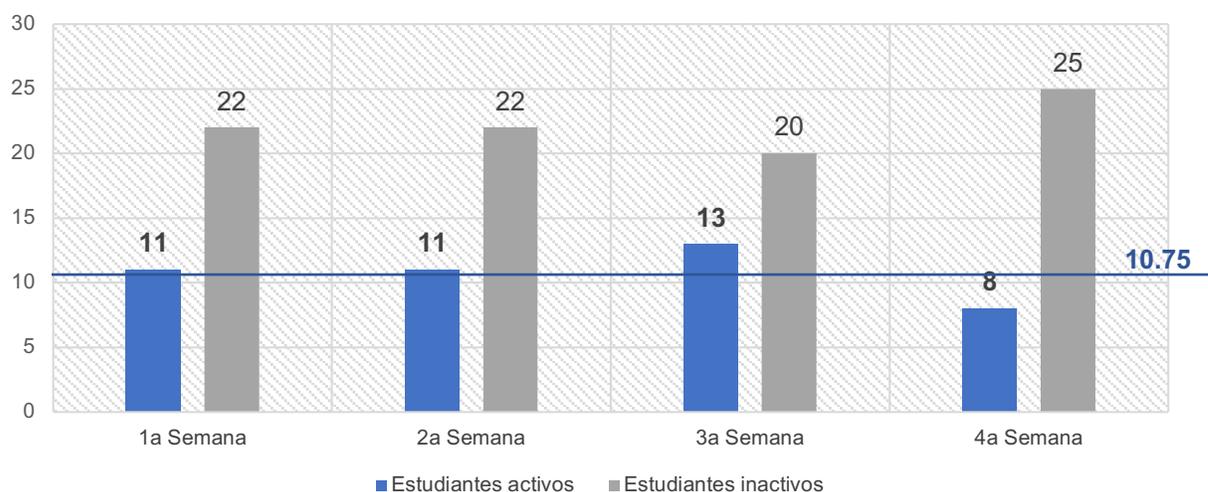
**Nota.** Las unidades de medida son: \*estudiantes y \*\*textos leídos o palabras (de acuerdo con la actividad).

En los siguientes apartados se desglosa la información contenida en la Tabla 17 con la finalidad de observar el progreso semanal de los estudiantes.

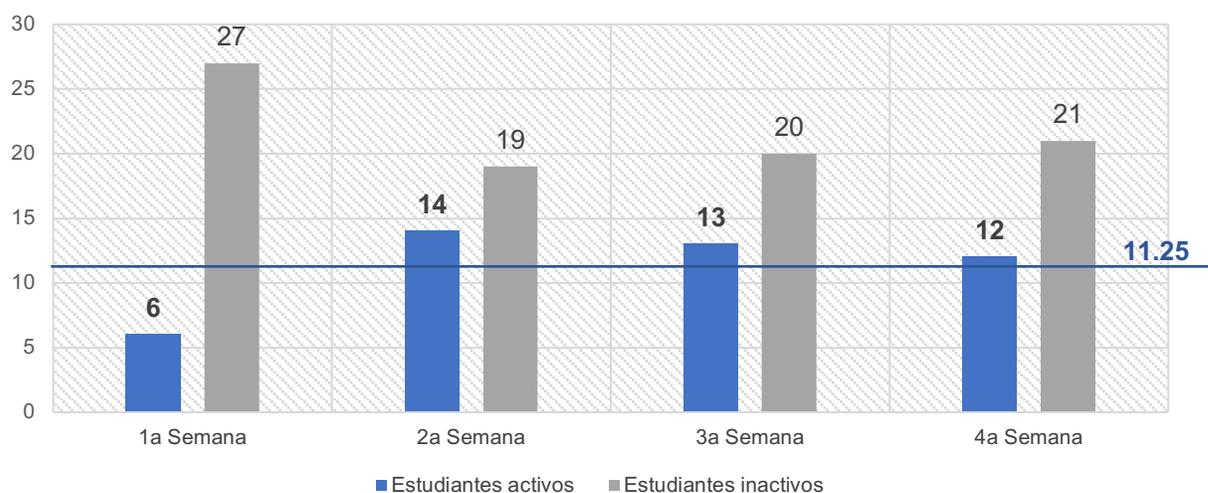
## 6.1 Datos muestreados

### 6.1.1 Textos leídos de la actividad lectura de textos

Se realizó el seguimiento de la participación en la actividad *lectura de textos* en ambos grupos de estudio durante las cuatro semanas de la evaluación. La Figura 53 muestra la participación que tuvo el grupo experimental en dicha actividad mientras que la Figura 54 presenta la participación del grupo de control. Se puede observar que, *en promedio, el grupo de control tuvo una mayor participación en la actividad lectura de textos con 11.25 estudiantes activos por semana en comparación con el grupo experimental que tuvo 10.75 estudiantes por semana.*



*Figura 53 Participación del grupo experimental en la actividad lectura de textos*

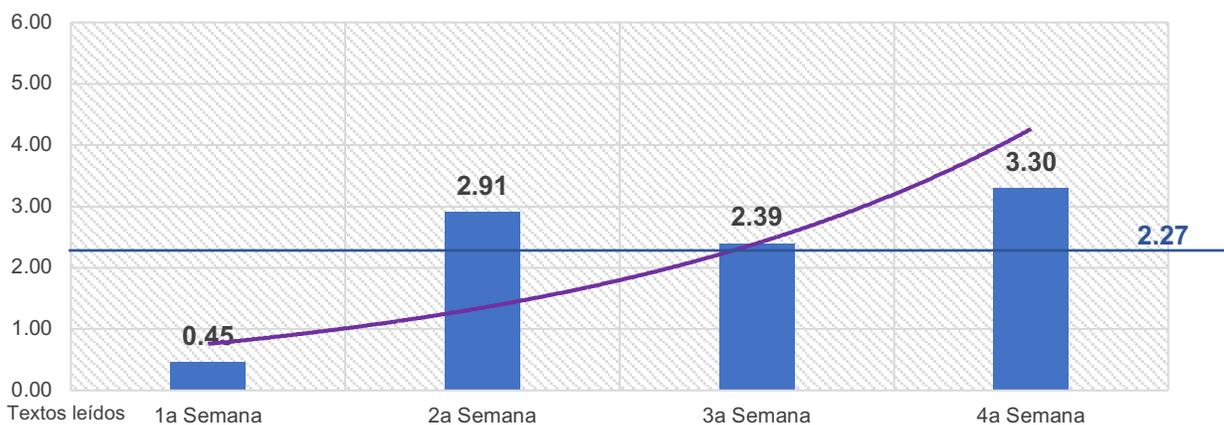


*Figura 54 Participación del grupo de control en la actividad lectura de textos*

Por otro lado, los grupos de estudio leyeron una cantidad reducida de textos por semana. La Figura 55 muestra la cantidad promedio de *textos leídos* por el grupo experimental durante cada semana de evaluación mientras que la Figura 56 presenta la cantidad promedio del grupo de control. Se observa que, *en promedio, el grupo experimental tuvo un mejor aprovechamiento con una cantidad de 2.36 textos leídos por semana en comparación con 2.27 textos leídos por semana del grupo de control.*



*Figura 55 Cantidad promedio de textos leídos por el grupo experimental*

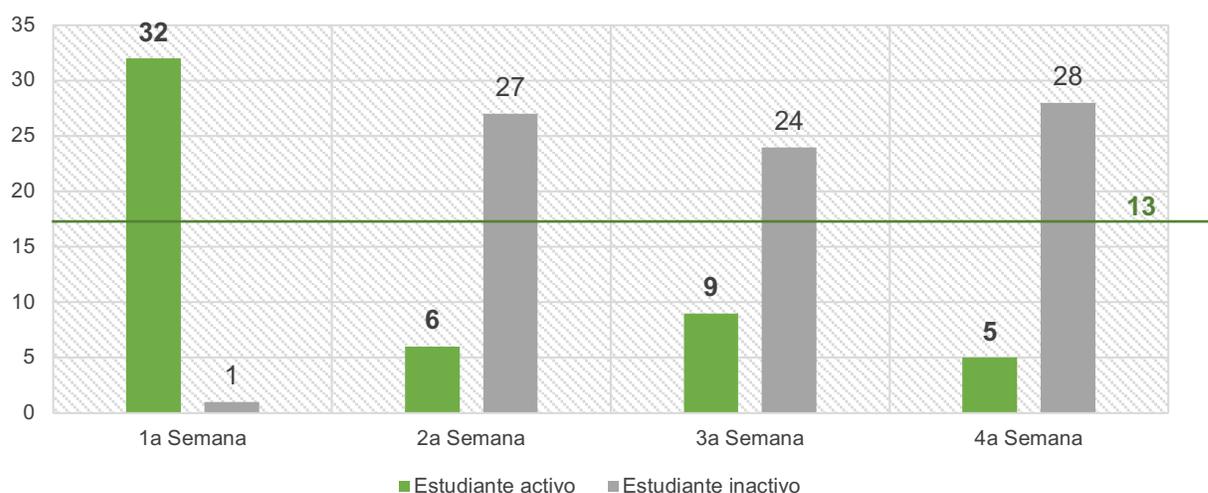


*Figura 56 Cantidad promedio de textos leídos por el grupo de control*

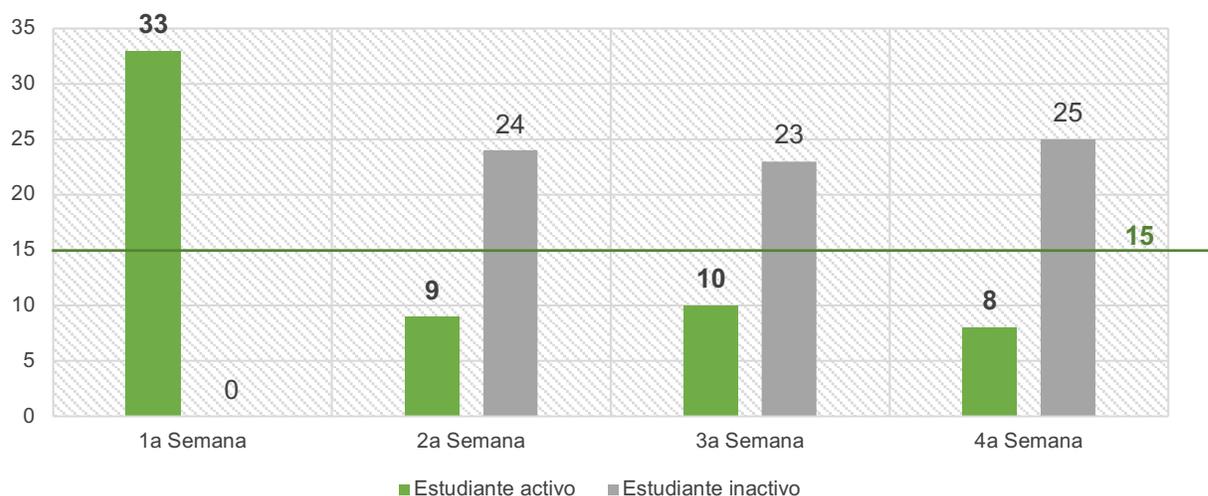
### **6.1.2 Palabras agregadas al vocabulario de la actividad ampliación de vocabulario**

Las *palabras agregadas al vocabulario* son palabras desconocidas que el estudiante agregó a su vocabulario de estudio durante una lectura, además, se toman en cuenta las palabras desconocidas que se agregaron automáticamente como resultado de la actividad de exploración realizada por el estudiante.

Se realizó el seguimiento de la participación en la actividad *ampliación de vocabulario* en ambos grupos de estudio durante las cuatro semanas de evaluación. La Figura 57 muestra la participación que tuvo el grupo experimental en dicha actividad mientras que la Figura 58 presenta la participación del grupo de control. Se puede observar que, *en promedio, el grupo de control tuvo una mayor participación en la actividad ampliación de vocabulario con 15 estudiantes activos por semana en comparación con el grupo experimental que tuvo 13 estudiantes por semana.*



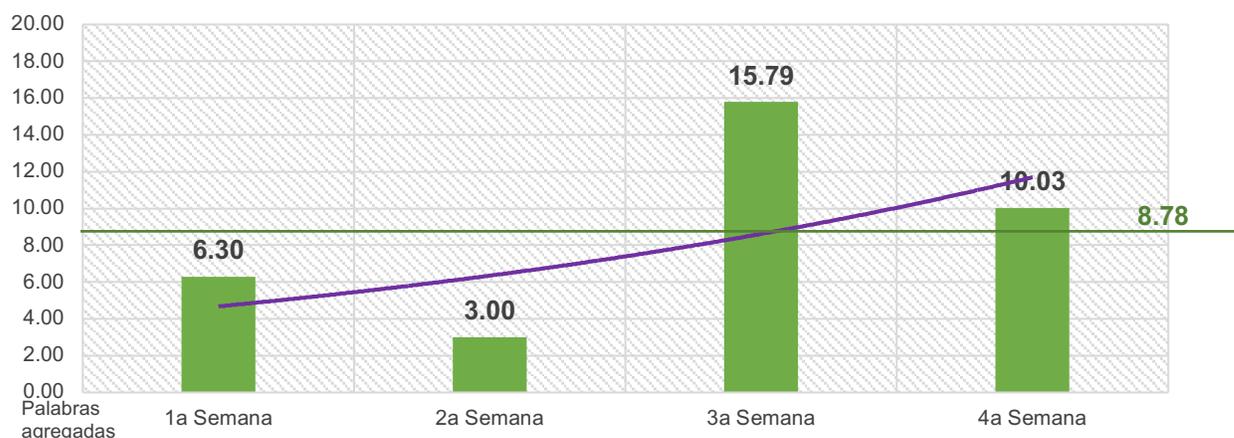
*Figura 57 Participación del grupo experimental en la actividad ampliación de vocabulario*



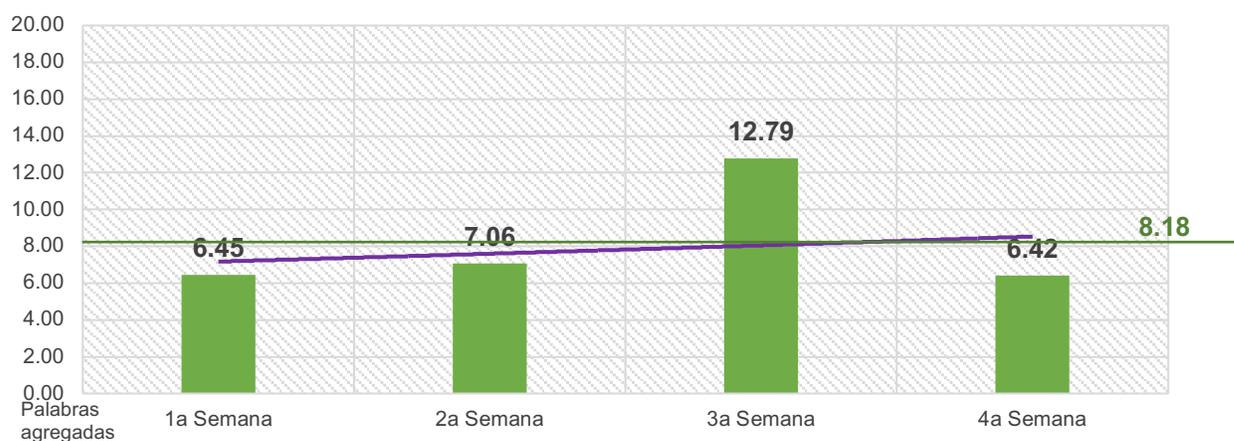
*Figura 58 Participación del grupo de control en la actividad ampliación de vocabulario*

También se analizó la cantidad de *palabras agregadas al vocabulario* de los grupos de estudio y se observó que fue una cantidad reducida de palabras agregadas por semana. La Figura

59 muestra la cantidad promedio de palabras agregadas al vocabulario por el grupo experimental durante cada semana de evaluación mientras que la Figura 60 presenta la cantidad promedio del grupo de control. Se observa que, *en promedio, el grupo experimental tuvo un mejor aprovechamiento con una cantidad de 8.78 palabras agregadas al vocabulario por semana en comparación con 8.18 palabras agregadas al vocabulario por semana del grupo de control.*



*Figura 59 Cantidad promedio de palabras agregadas al vocabulario por el grupo experimental*



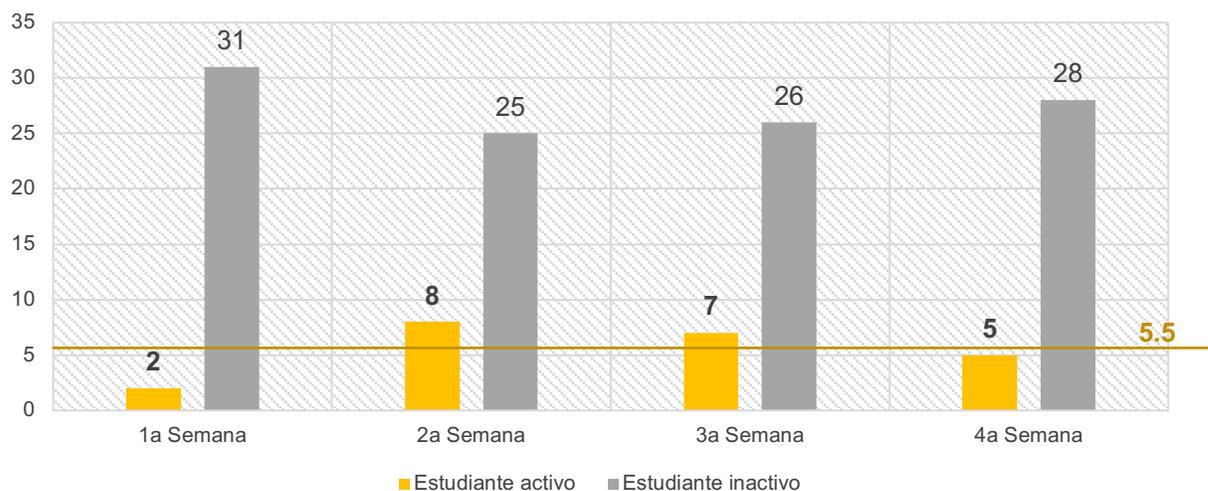
*Figura 60 Cantidad promedio de palabras agregadas al vocabulario por el grupo de control*

### **6.1.3 Palabras estudiadas de la actividad repaso de vocabulario**

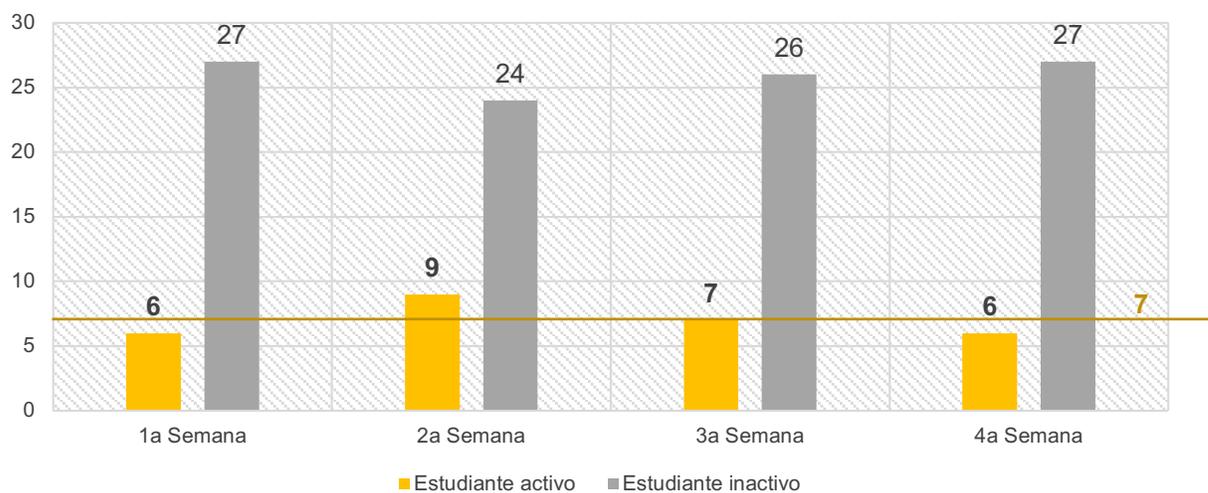
Las *palabras estudiadas* son las palabras que el estudiante ha repasado mediante la dinámica de tarjetas de memoria del sistema SAIC.

Se realizó el seguimiento de la participación del grupo experimental y de control en el estudio de palabras de la actividad repaso de vocabulario durante las cuatro semanas de evaluación. La Figura 61 muestra la participación que tuvo el grupo experimental en dicha actividad mientras

que la Figura 62 presenta la participación del grupo de control. Se puede observar que, *en promedio, el grupo de control tuvo una mayor participación estudiando palabras con 7 estudiantes activos por semana en comparación con el grupo experimental que tuvo 5.5 estudiantes por semana.*



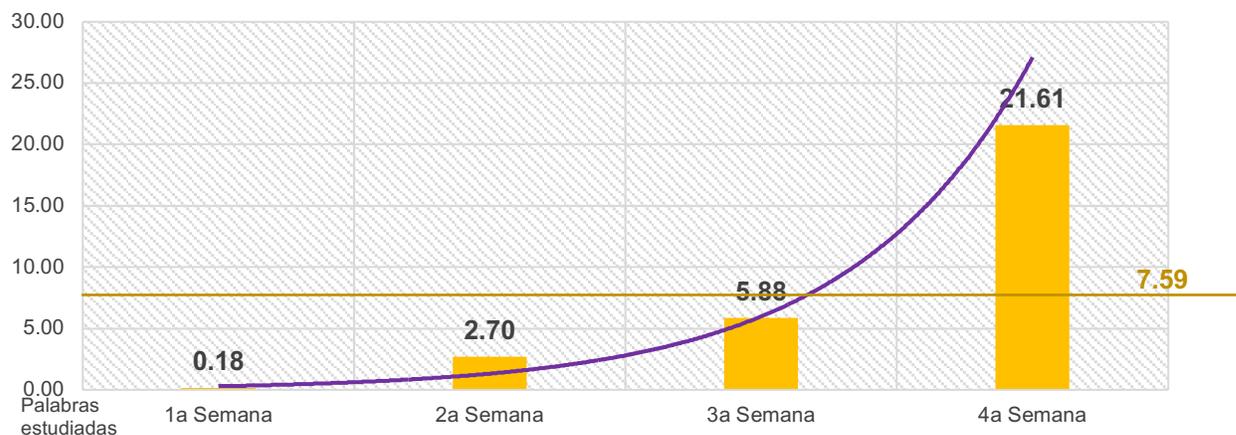
*Figura 61 Participación del grupo experimental estudiando palabras de la actividad repaso de vocabulario*



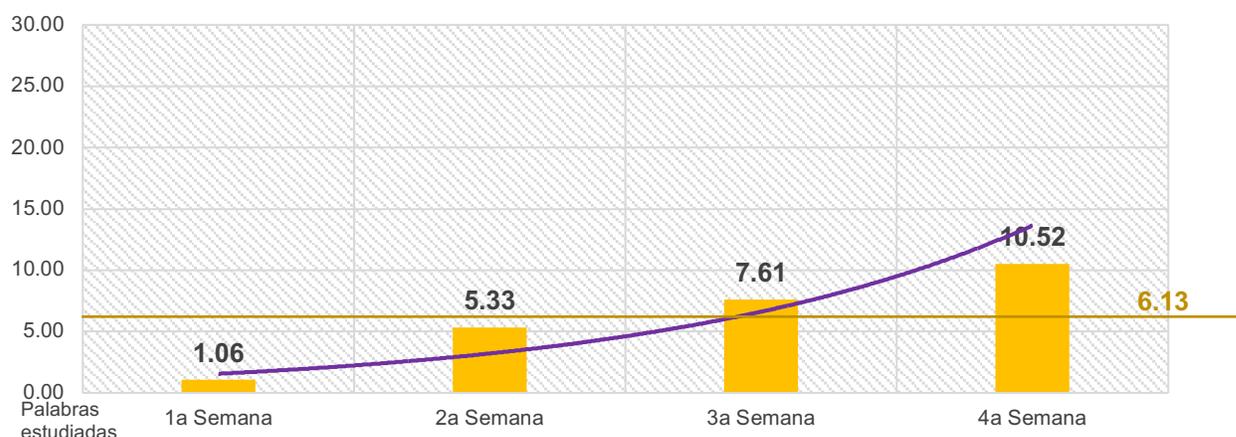
*Figura 62 Participación del grupo de control estudiando palabras de la actividad repaso de vocabulario*

También se analizó la cantidad de *palabras estudiadas* de los grupos de estudio y se observó que los estudiantes repasaron una cantidad reducida de palabras por semana. La Figura 63 muestra la cantidad promedio de palabras estudiadas por el grupo experimental durante cada semana de evaluación mientras que la Figura 64 presenta la cantidad promedio del grupo de control. Se observa que, *en promedio, el grupo experimental tuvo un mejor aprovechamiento*

con una cantidad de 7.59 palabras estudiadas por semana en comparación con 6.13 palabras estudiadas por semana del grupo de control.



*Figura 63 Cantidad promedio de palabras estudiadas por el grupo experimental*



*Figura 64 Cantidad promedio de palabras estudiadas por el grupo de control*

#### **6.1.4 Palabras en proceso de aprendizaje de la actividad repaso de vocabulario**

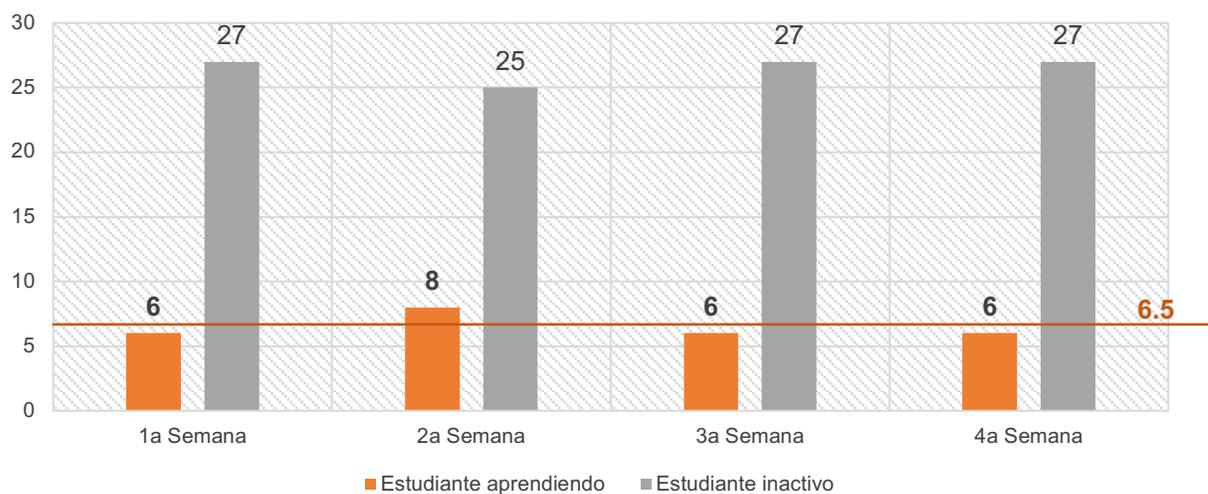
Las *palabras en proceso de aprendizaje* o *palabras en aprendizaje* son el resultado de las palabras estudiadas de la actividad repaso de vocabulario. Las palabras en aprendizaje representan un indicador sobre las palabras que los estudiantes están aprendiendo dentro del sistema SAIC.

Se analizaron ambos grupos de estudio con la finalidad de obtener qué participación de estudiantes está aprendiendo palabras durante las cuatro semanas de evaluación. La Figura 65 muestra la cantidad promedio de estudiantes del grupo experimental que lograron palabras en proceso de aprendizaje mientras que la Figura 66 presenta la cantidad promedio del grupo de

control. Se puede observar que, *en promedio, el grupo de control tuvo una mayor participación aprendiendo palabras con 6.5 estudiantes activos por semana en comparación con el grupo experimental que tuvo 5 estudiantes por semana.*



*Figura 65 Estudiantes del grupo experimental que lograron palabras en proceso de aprendizaje*



*Figura 66 Estudiantes del grupo de control que lograron palabras en proceso de aprendizaje*

Por otro lado, se analizó la cantidad de *palabras en proceso de aprendizaje* de los grupos de estudio y se observó que los estudiantes están aprendiendo una cantidad reducida de palabras por semana. La Figura 67 muestra la cantidad promedio de palabras en proceso de aprendizaje que lograron los estudiantes del grupo experimental durante cada semana de evaluación mientras que la Figura 68 presenta la cantidad promedio para el grupo de control. Se observa que, *en promedio, el grupo experimental tuvo un mejor aprovechamiento con una cantidad de*

6.64 palabras en proceso de aprendizaje por semana en comparación con 4.17 palabras en proceso de aprendizaje por semana del grupo de control.

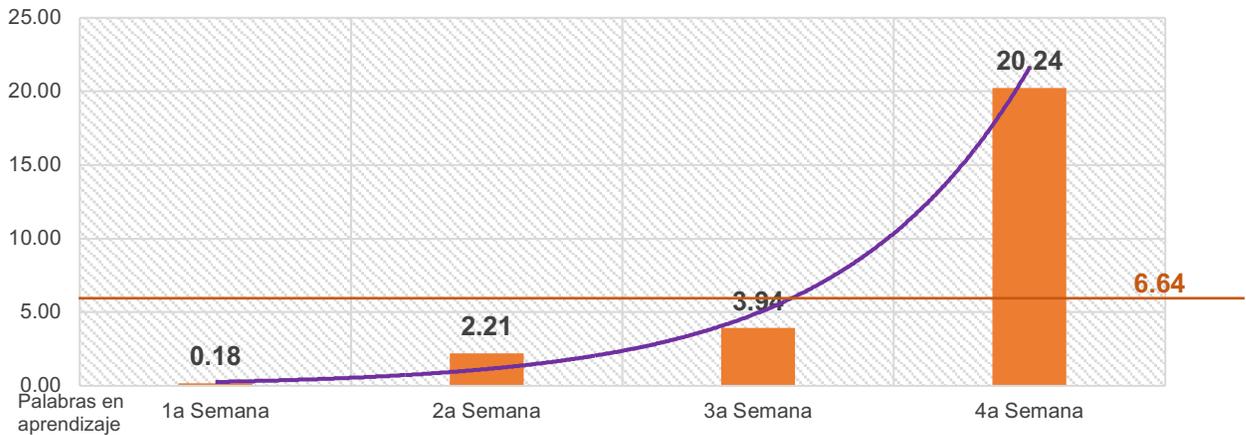


Figura 67 Cantidad promedio de palabras en proceso de aprendizaje por el grupo experimental

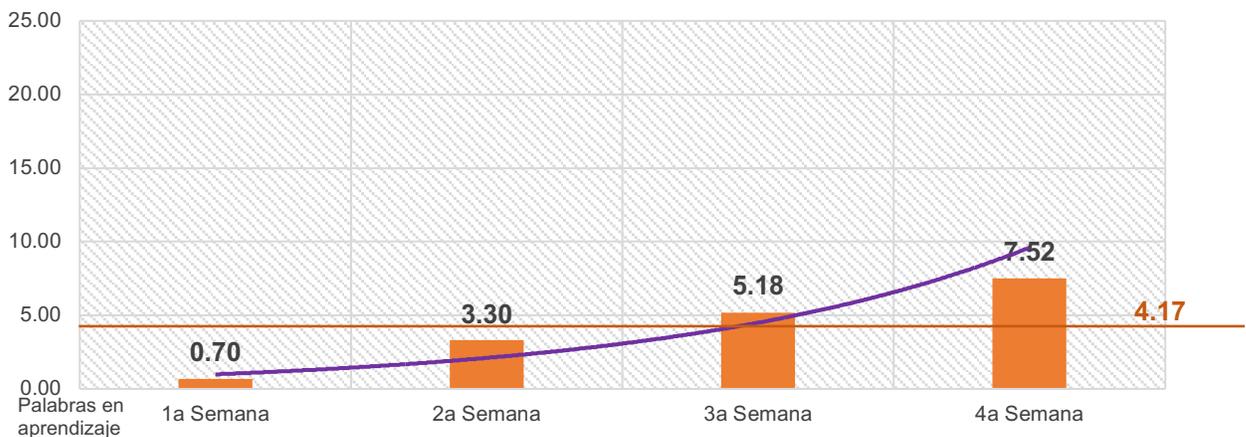


Figura 68 Cantidad promedio de palabras en proceso de aprendizaje por el grupo de control

## 6.2 Medida pretest-postest

En la evaluación, se tomó en cuenta que los estudiantes participantes ya pertenecían a grupos existentes de una población de estudiantes de nivel medio superior (Colegio de Bachilleres), por tal motivo, se redistribuyeron en grupo experimental y de control en el momento de su registro mediante una asignación aleatoria simple, es decir, los estudiantes se asignaron por medio de un algoritmo de aleatoriedad al momento de su registro, cabe mencionar que no se contaba con un orden de lista ni una fecha límite de registro.

La medida *pretest-postest* se determinó por medio de sus dos magnitudes: *pretest*, se obtuvo la diferencia de conocimientos entre el grupo experimental y el grupo de control antes que fueran

expuestos a las actividades de aprendizaje del sistema SAIC (diferencia de conocimientos antes del tratamiento), estos conocimientos se calcularon mediante la actividad de exploración; y *postest*, se obtuvo la diferencia de conocimientos entre el grupo experimental y el grupo de control después que fueran expuestos a las actividades de aprendizaje del sistema SAIC (diferencia de conocimientos después del tratamiento), estos conocimientos se atribuyen al estudio durante la evaluación. La Figura 69 expone la resolución de la medida *pretest-postest* para los datos muestreados: *actividad de exploración* para la diferencia de conocimientos antes del tratamiento y *palabras en proceso de aprendizaje* para la diferencia de conocimientos después del tratamiento.

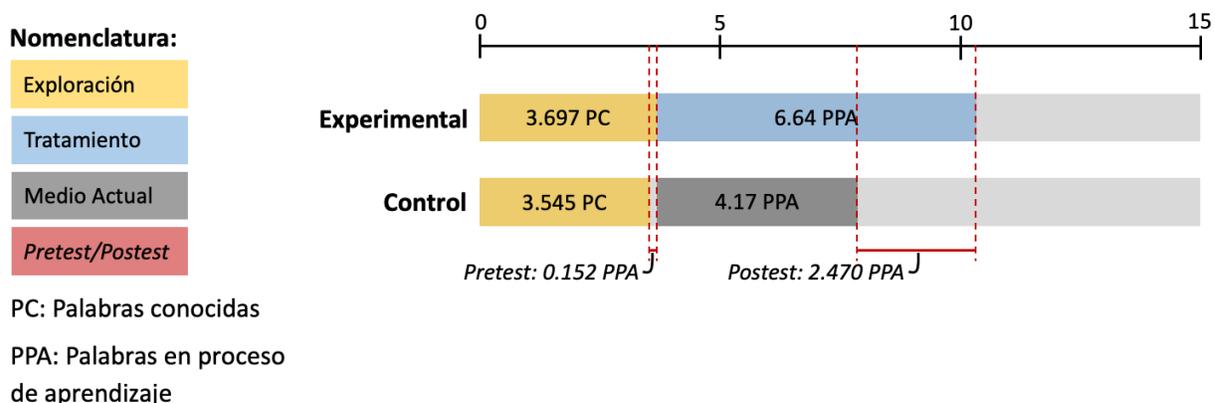


Figura 69 Resolución de la medida *pretest-postest*

De acuerdo con la literatura, si *postest* es mayor que *pretest*, entonces el tratamiento surtió efecto, como se observa en la Figura 69. Esto significa que *los estudiantes del grupo experimental que utilizaron el sistema SAIC con todas las características descritas en el método del presente trabajo aportó que tuvieron un mayor aprendizaje de palabras en inglés en comparación con los estudiantes del grupo de control que utilizaron el sistema SAIC con restricciones* (simulando las funciones ofrecidas por el sitio *readlang.com*, representado en la Figura 69 como *medio actual*).

### 6.3 Prueba de hipótesis de la diferencia de medias

La evaluación proporcionó dos medias que representan el aprendizaje de palabras en inglés para el grupo experimental y el grupo de control, con base en estas medias y mediante la medida *pretest-postest*, se concluyó que el aprendizaje del grupo experimental fue mayor en comparación con el obtenido por el grupo de control; sin embargo, este análisis no indica si los

resultados obtenidos se deben al azar o si en realidad son dos muestras provenientes de poblaciones con medias distintas.

La *prueba de hipótesis de la diferencia de medias* se aplicó para comprobar si la diferencia observada entre las medias muestrales es estadísticamente significativa, es decir, si la diferencia en el aprendizaje de palabras en inglés se atribuye al tratamiento aplicado. Las muestras obtenidas en los resultados de la evaluación no corresponden a una distribución normal, no obstante, la literatura sugiere realizar una prueba paramétrica asumiendo una distribución normal cuando la cantidad de datos es mayor o igual a 30. A continuación, se presenta la información y componentes que se utilizaron en la prueba de hipótesis.

- La hipótesis nula ( $H_0$ ) y la hipótesis alternativa ( $H_1$ ) que se muestran en la Figura 70.

**$H_0$ :  $\mu_1 = \mu_2$** , la media de palabras en proceso de aprendizaje (PPA) de la población con tratamiento que ha sido muestreada (grupo experimental) es igual a la media de palabras en proceso de aprendizaje (PPA) de la población sin tratamiento que ha sido muestreada (grupo de control).

**$H_1$ :  $\mu_1 > \mu_2$** , la media de palabras en proceso de aprendizaje (PPA) de la población con tratamiento que ha sido muestreada (grupo experimental) es mayor a la media de palabras en proceso de aprendizaje (PPA) de la población sin tratamiento que ha sido muestreada (grupo de control).

*Figura 70 Hipótesis nula y alternativa*

Donde:  $\mu_1$  es la media de palabras en proceso de aprendizaje (PPA) de la población con tratamiento que ha sido muestreada (grupo experimental) y  $\mu_2$  es la media de palabras en proceso de aprendizaje (PPA) de la población sin tratamiento que ha sido muestreada (grupo de control).

- Los resultados obtenidos de la evaluación para las palabras en proceso de aprendizaje (PPA) expuestos en la Tabla 18.

*Tabla 18 Información sobre las palabras en proceso de aprendizaje (PPA)*

Variable	Valores		Unidades
	Grupo experimental	Grupo de control	
Cantidad de datos ( $n$ )	33	33	Elementos
Media ( $\bar{x}$ )	6.644	4.174	PPA
Desviación estándar ( $s$ )	20.654	12.367	PPA
Varianza ( $s^2$ )	426.586	157.947	PPA <sup>2</sup>

- La fórmula para  $z$  calculado que se muestra en la ecuación (8) y que tiene su origen en la *fórmula del valor normal* o *valor z* que se presenta en la ecuación (9). Los puntajes  $z$  son transformaciones que se pueden convertir en valores de una distribución normal, con el propósito de analizar su distancia respecto a la media (esta distancia se expresa en unidades de desviación estándar). En este caso, se desconocía la desviación estándar de ambas poblaciones, por lo tanto, se utilizó la desviación estándar muestral con base en que esta desviación es una aproximación a la desviación estándar poblacional.

$$z_c = \frac{(\bar{x}_1 - \bar{x}_2) - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (8)$$

Donde:  $z_c$  es el valor  $z$  calculado,  $\bar{x}_1$  es la media de PPA del grupo experimental,  $\bar{x}_2$  es la media de PPA del grupo de control,  $\delta$  es la diferencia esperada entre medias,  $\sigma_1^2$  es la varianza poblacional del grupo experimental,  $\sigma_2^2$  es la varianza poblacional del grupo de control,  $n_1$  es la cantidad de datos de la muestra del grupo experimental y  $n_2$  es la cantidad de datos de la muestra del grupo de control.

$$z = \frac{x - \mu}{\sigma} \quad (9)$$

Donde:  $z$  es un valor normal o valor  $z$  de la tabla de valores de probabilidad acumulada para la distribución normal estándar,  $x$  es el valor de la muestra,  $\mu$  es el valor esperado y  $\sigma$  es la desviación estándar de los datos.

- La *tabla de valores de probabilidad acumulada para la distribución normal estándar* que se utilizó para calcular  $z$  tabular, esta tabla se presenta en la Figura 71.

z	0	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505

Figura 71 Tabla de valores de probabilidad acumulada para la distribución normal estándar (Vargas, 2010)

La prueba de hipótesis de la diferencia de medias se realizó con un intervalo de confianza del 95% (o un nivel de significancia del 5%). El *nivel de significancia* es la probabilidad de rechazar la hipótesis nula cuando es verdadera, también llamada error de tipo 1; por el contrario, el error de tipo 2, se comete si se acepta la hipótesis nula cuando es falsa. La prueba de hipótesis tuvo como resultado un valor  $z_c = 0.589$  y un valor  $z_t = 1.645$ , por lo tanto,  $z_c < z_t$ .

La Figura 72 muestra  $z_c$  y  $z_t$  en la distribución normal con un intervalo de confianza del 95% con una cola. Se utiliza una cola para el intervalo debido a que se desea corroborar estadísticamente la conclusión que dice que *el aprendizaje de palabras en inglés del grupo experimental es mayor en comparación con el aprendizaje obtenido por el grupo de control*, por lo tanto, se observa que cualquier valor por encima de  $z_t$  satisface esta conclusión.

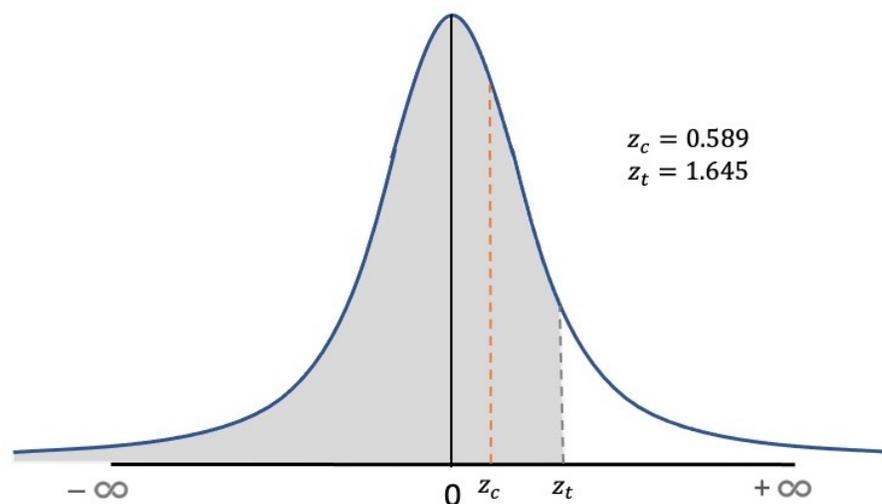


Figura 72 Gráfica de la posición de  $z_c$  y  $z_t$  en la curva de distribución normal estándar

Con base en lo anterior, se acepta la hipótesis nula ( $H_0$ ) y se puede concluir que *la diferencia entre la media de palabras en proceso de aprendizaje (PPA) de la población con tratamiento que se ha muestreado (grupo experimental) y la media de palabras en proceso de aprendizaje (PPA) de la población sin tratamiento que se ha muestreado (grupo de control) se atribuye al azar y no al tratamiento aplicado de acuerdo con la evidencia estadística analizada*. En otras palabras, el resultado que indica que el grupo experimental presenta un mayor aprovechamiento en las palabras en proceso de aprendizaje (PPA) en comparación con el grupo de control no es estadísticamente significativo.

#### **6.4 Discusión**

En la prueba de hipótesis realizada se concluyó que los resultados obtenidos en la evaluación no son estadísticamente significativos, es decir, no existe una diferencia estadística entre el aprovechamiento en el aprendizaje de palabras en inglés del grupo experimental y el aprovechamiento del grupo de control. Esta aseveración podría suponer que las estrategias propuestas en el método del presente trabajo no surtieron efecto, sin embargo, el resultado de la prueba de hipótesis se puede deber a la poca participación de los estudiantes en la actividad *repaso de vocabulario*.

La Figura 73 presenta un gráfico histórico sobre las *palabras en proceso de aprendizaje* de la actividad *repaso de vocabulario* del grupo experimental mientras que en la Figura 74 se muestra la misma información para el grupo de control. En estas figuras se puede observar el promedio de palabras en proceso de aprendizaje (PPA) de ambos grupos de estudio durante la evaluación, estos promedios exponen la poca e inconsistente participación que dichos grupos han tenido. Cabe resaltar que algunos estudiantes presentan participaciones elevadas que en las figuras se aprecian como picos en la gráfica, estos picos de participación provocaron un aumento en la dispersión de los datos con respecto a su media (desviación estándar más amplia) para cada grupo de estudio.

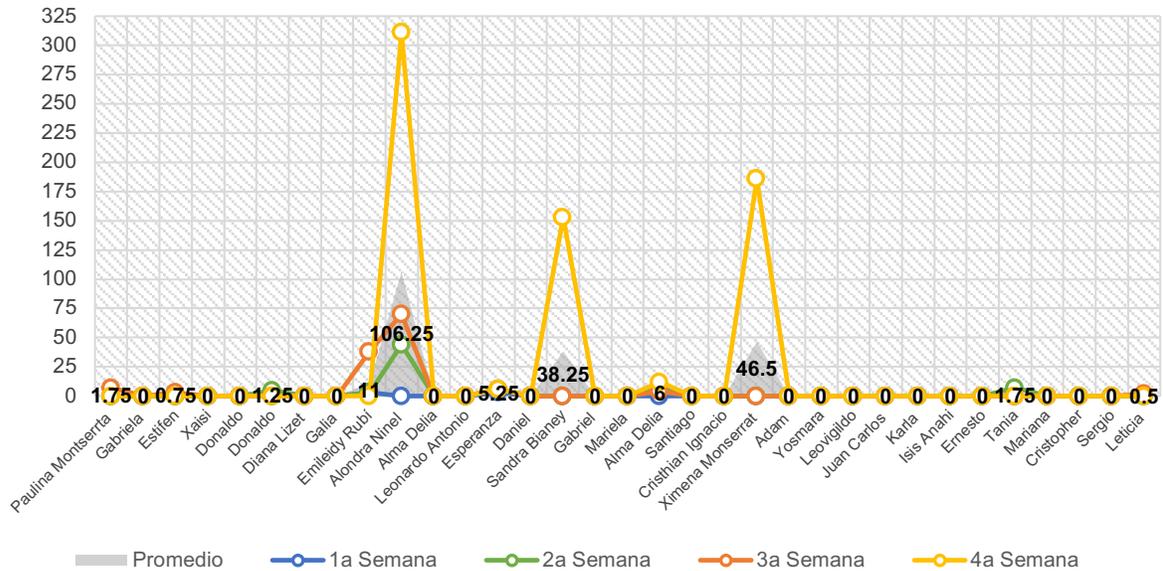


Figura 73 Histórico de palabras en proceso de aprendizaje del grupo experimental

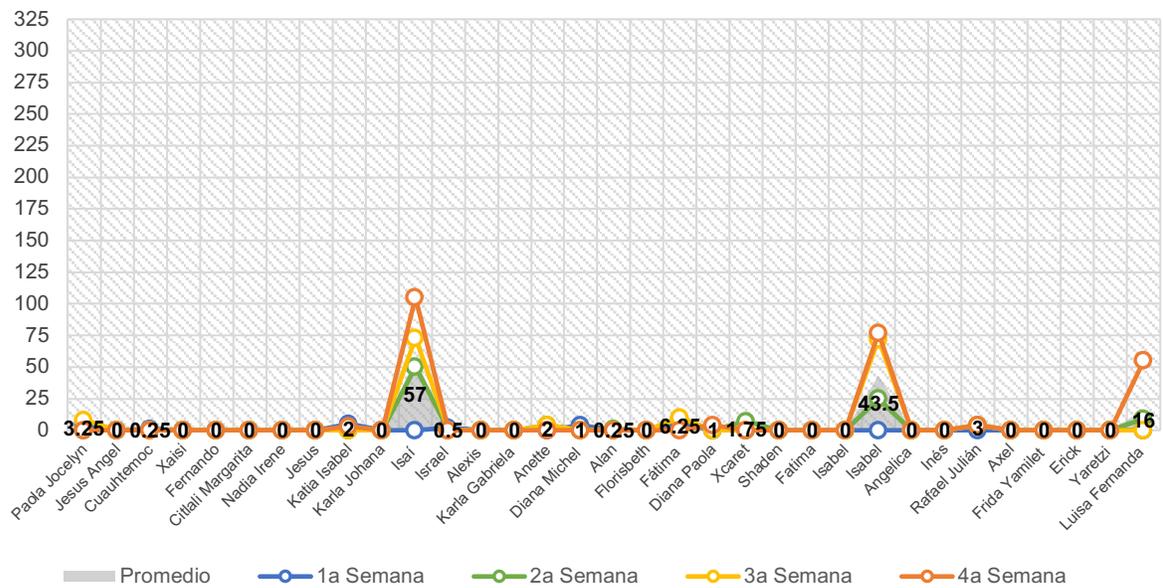


Figura 74 Histórico de palabras en proceso de aprendizaje del grupo de control

Por otro lado, la prueba de hipótesis aplicada en el presente trabajo es sensible a la desviación estándar de la diferencia de las medias de los grupos de estudio como se puede apreciar en la ecuación (9) utilizada en dicha prueba.

$$Z = \frac{x - \mu}{\sigma} \quad (9)$$

Donde:  $z$  es un valor normal o valor  $z$  de la tabla de valores de probabilidad acumulada para la distribución normal estándar,  $x$  es el valor de la muestra,  $\mu$  es el valor esperado y  $\sigma$  es la desviación estándar de los datos.

Por lo tanto, si la dispersión de los datos o desviación estándar es mayor a la diferencia entre las medias muestrales, como en los resultados de la evaluación; entonces la diferencia entre medias no es significativa. Con base en esto, el resultado se atribuye a la poca e inconsistente participación de la mayoría de los estudiantes y a los picos de participación de unos cuantos estudiantes.

# Capítulo 7 Conclusiones

Con esta investigación se logró crear un método que unifica la investigación, diseño, desarrollo, implementación y evaluación de un sistema de software con la finalidad de apoyar a los estudiantes de una lengua extranjera con el aprendizaje de vocabulario mediante la asistencia computacional. Es importante mencionar que el método implementado propone la mejora continua del enfoque de enseñanza o la experiencia de usuario con el software.

También se tiene como resultado de la investigación un repositorio con el sistema SAIC para el aprendizaje de vocabulario basado en la lectura de textos cortos en inglés, recomendación de textos y el estudio de palabras desconocidas por medio de tarjetas de memoria con la técnica de repetición espaciada. Además, se tiene un repositorio con algunas herramientas de software para conexión de bases de datos, procesamiento de textos, recomendación de textos, repetición espaciada, conexión de servicios de traducción, búsqueda de imágenes y texto a voz con el lenguaje de programación Ruby. Si se desea más información al respecto de estos repositorios, se puede dirigir al correo electrónico [tecn@cenidet.tecnm.mx](mailto:tecn@cenidet.tecnm.mx).

Por otro lado, se tuvo que limitar la evaluación a cuatro semanas por cuestiones de tiempo, esto fue derivado al confinamiento por la pandemia de covid19 que retrasó y dificultó la aplicación y seguimiento de las pruebas a los grupos de estudio. En consecuencia, no fue posible realizar la evaluación de palabras aprendidas (que no son igual a las palabras en proceso de aprendizaje que se han analizado) como lo sugiere la teoría de la curva del olvido mediante el número máximo de repeticiones espaciadas durante el periodo de evaluación.

Por último, la prueba de hipótesis de la diferencia de medias indicó que el resultado obtenido en la evaluación no es estadísticamente significativo, una de las razones que pudieron provocar este resultado se debe a la poca e inconsistente participación de los estudiantes en las actividades del estudio. Sin embargo, se tiene la investigación teórica necesaria para que en un trabajo futuro se pueda realizar la evaluación de palabras aprendidas como lo sugiere la teoría de la curva del olvido si se obtiene la participación suficiente de los grupos de estudio.

## 7.1 Mejoras técnicas y funcionales para el sistema SAIC

Durante la evaluación, se detectaron áreas de mejora técnica y funcional debido a la puesta en producción del sistema SAIC, el uso que le dieron los estudiantes participantes y los comentarios realizados por la docente a cargo de los grupos de estudio. A continuación, se expone una lista de mejoras clasificada por categorías:

- General.
  - Adquirir un certificado de seguridad para utilizar conexiones HTTPS.
  - Adquirir un dominio para facilitar el acceso al sistema SAIC.
  - Proporcionar un mecanismo que proporcione los códigos de recuperación de contraseña del sistema SAIC mediante correo electrónico.
  - Proporcionar la función de administración de sesiones abiertas y cierre de sesión remoto.
  - Implementar el cierre de sesión automático después de cierto tiempo de inactividad.
  - Seguir la ubicación del usuario en las páginas del sistema SAIC para mejorar la navegación entre páginas.
  - Implementar nuevas actividades para el aprendizaje de los estudiantes, esta implementación se debe centrar en actividades activas de acuerdo con el cono de la experiencia de Dale (1932).
- Base de datos.
  - Almacenar la imagen y miniatura como binario en la base de datos para evitar la pérdida de información en caso de que la imagen ya no esté disponible en su dirección URL original.
  - Investigar la posibilidad de almacenar el audio de la pronunciación de palabras como binario en la base de datos.
- Algoritmo de repetición espaciada.
  - Asignar la retención o probabilidad de recordar la información de estudio ( $p$ ) de manera dinámica de acuerdo con el momento de repaso del estudiante.

- Solicitar al estudiante su hora favorita del día para el estudio de palabras desconocidas y calcular  $p$  de acuerdo con dicha hora.
- Lista de lecturas.
  - Implementar la clasificación de textos por categorías.
  - Implementar la búsqueda de lecturas por título.
  - Implementar el ordenamiento de lecturas por fecha de creación y alfabéticamente.
- Lecturas.
  - Implementar un visor de lectura compatible con funciones avanzadas de selección.
  - Implementar un editor de lectura compatible con funciones avanzadas de formato y edición.
  - Proporcionar la posibilidad de compartir lecturas con otros estudiantes.
  - Marcar palabras que el usuario ya tiene en su vocabulario.
  - Mostrar varias opciones de traducción de una palabra.
  - Implementar la traducción por oración seleccionada.
  - Desarrollar la opción para agregar una palabra al vocabulario automáticamente cuando se traduzca y un algoritmo que seleccione la mejor imagen para esa palabra de forma automática.
  - Proporcionar el conteo y limitación de palabras (a una cuartilla) desde el editor de lecturas.
  - Permitir introducir, mostrar y editar una referencia de autor para el texto.
  - Gestionar la privacidad de la lectura (pública, grupo, privada).
  - Proporcionar la función de marcar una lectura como leída o no leída.
- Vocabulario.
  - Proporcionar la posibilidad de agregar palabra al vocabulario de forma manual, sin necesidad de hacerlo desde una lectura.

- Editar palabra existente en el vocabulario.
- Eliminar palabra existente en el vocabulario.
- Usuario.
  - Agregar al registro el nivel educativo del estudiante con fines estadísticos y de mejora del método.
  - Proporcionar al usuario de tipo *estudiante* un reporte estadístico de su cuenta.
  - Implementar un asistente de estudio (apartado o asistente visual que guíe al estudiante en sus actividades diarias dentro del sistema).
  - Agregar el usuario de tipo *administrador*, además, los módulos necesarios para administrar los usuarios, lecturas, palabras y grupos del sistema SAIC.
- Grupos.
  - Agregar el usuario de tipo *docente*.
  - Implementar un módulo para administrar grupos de estudiantes, accesible sólo por el usuario de tipo *docente*.
  - Implementar un reporte estadístico sobre los grupos que administra un usuario de tipo *docente*.

## 7.2 Trabajos futuros

Durante la investigación, se analizaron distintos enfoques de aprendizaje que pueden ser objeto de investigación para reorientar el enfoque propuesto por el método del presente trabajo, con la finalidad de mejorar y agregar nuevas características al sistema SAIC.

Se recomienda realizar un estudio longitudinal de más de un mes, con la finalidad de obtener la participación de los grupos de estudio suficiente para realizar la evaluación de palabras aprendidas como lo sugiere la teoría de la curva del olvido expuesta en este trabajo. También se sugiere una evaluación de la experiencia de usuario para mejorar las funciones del sistema y el seguimiento de actividades, con el objetivo que los estudiantes realicen las actividades de estudio en tiempo y forma. Además, se propone una lista de mejoras clasificada por categorías que se presenta en el apartado *mejoras técnicas y funcionales para el sistema SAIC*.

Por último, se sugiere una evaluación con estudiantes de distinto nivel académico para poder observar si hay una mejora en la participación o en el aprovechamiento en comparación con los grupos de nivel medio superior.

# Referencias

- Agencia Mexicana de Cooperación Internacional para el Desarrollo. (2017). *Programa de Capacitación de Estudiantes SEP-SRE PROYECTA 100,000 | Estados Unidos de América 2017*. Recuperado el 12 de noviembre de 2018, de Agencia Mexicana de Cooperación Internacional para el Desarrollo: <https://www.gob.mx/amexcid/documentos/programa-de-capacitacion-de-estudiantes-sep-sre-proyecto-100-000-estados-unidos-de-america-2017>
- Al-Shumari, M., & Bella, G. (2014). Online english vocabulary learning on different systems for non-english speakers. *Proceedings ELMAR-2014*.
- Allahyari, M., Pouriye, S., Assefi, M., Safaei, S., Trippe, E., Gutierrez, J., & Kochut, K. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *In Proceedings of KDD Bigdas*.
- Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, 21-34.
- Ausubel, D. P., & Youssef, M. (1965). The effect of spaced repetition on meaningful retention. *Journal of General Psychology*, 1(73), 147-150.
- Bärnighausen, T., Röttingen, J.-A., Rockers, P., Shemilt, I., & Tugwell, P. (2017). Quasi-experimental study designs series – paper 1: history and introduction. *Journal of Clinical Epidemiology*.
- Beccari, A., & Penavel, C. (2014). El enfoque inductivo-contextual de "La Lengua Latina ilustrada por sí misma" por Hans Henning Orberg. *PhaoS*(14), 123-142.
- Bleier, S. (2010). *NLTK's list of english stopwords*. Obtenido de GitHub: <https://gist.github.com/sebleier/554280>
- British Council. (2021). *Learn English Beginner A1*. Recuperado el Noviembre de 2021, de British Council Learn English: <https://learnenglish.britishcouncil.org/skills/reading/beginner-a1>
- Cambridge Assessment English. (2020). *Test Your English*. Obtenido de Cambridge Assessment English: <https://www.cambridgeenglish.org/test-your-english/>
- Chalermbuntai, R., Kittipol, R., Ranong, M., & Tangworakitthaworn, P. (2017). The Design and Development of the Vocabulary Learning System for Second Language Learners using Word Association. *2017 6th ICT International Student Project Conference (ICT-ISPC)*.
- Chmiel, A., & Schubert, E. (2018). Using Psychological Principles of Memory Storage and Preference to Improve Music Recommendation Systems. *Leonardo Music Journal*, 28, 77-81.
- Chopra, A., Prashar, A., & Sain, C. (2013). Natural Language Processing. *International journal of technology enhancements and emerging engineering research*, 1, 131-134.
- Consejo de Europa. (2001). *Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Recuperado el 11 de Septiembre de 2020, de Consejo de Europa: <https://rm.coe.int/1680459f97>
- Dale, E. (1932). Methods for Analyzing the Content of Motion Pictures. *Journal of Educational Sociology*, 6, 244-250.
- De Saint-Exupéry, A. (2001). *The Little Prince*. Continental Book Company.

- Ebbinghaus, H. (1885). *Memory: A Contribution to Experimental Psychology*. New York: Dover.
- EF Education First. (2020). *EF EPI 2020*. Recuperado el 12 de noviembre de 2020, de EF Education First: <https://www.ef.com.mx/epi/downloads/>
- Ferreira, A., Vine, A., & Elejalde, J. (2015). Diseño y evaluación de una aplicación tecnológica para la enseñanza del español como lengua extranjera. *ONOMÁZEIN: Revista semestral de lingüística, filología y traducción*, 145-166.
- Flores, E., & Maureira, F. (2015). Estilos de aprendizaje V.A.K. en estudiantes de Educación Física y otras pedagogías en la Universidad Internacional SEK de Chile. *Viref: Revista de Educación Física*, 4(2), 14-24.
- Garner, J. (2021). *VocabKitchen*. Recuperado el Diciembre de 2021, de VocabKitchen: <https://www.vocabkitchen.com/home>
- Gobierno de México. (2011). *NORMA Oficial Mexicana NOM-003-SEGOB-2011*. Recuperado el Septiembre de 2021, de Gobierno de México: <https://www.gob.mx/cms/uploads/attachment/file/138413/NOM-003-SEGOB-2011.pdf>
- Google Inc. (2021). *Cloud Text To Speech*. Recuperado el Octubre de 2021, de Google Cloud: <https://cloud.google.com/text-to-speech/>
- Google Inc. (2021). *Cloud Translation*. Recuperado el Octubre de 2021, de Google Cloud: <https://cloud.google.com/translate/>
- Google Inc. (2021). *Programmable Search Engine*. Recuperado el Octubre de 2021, de Google: <https://programmablesearchengine.google.com/about/>
- GoRails, LLC. (2021). *Deploy Ruby On Rails*. Recuperado el Noviembre de 2021, de GoRails: <https://gorails.com/deploy/ubuntu/20.04>
- Gribbons, B., & Herman, J. (1996). True and Quasi-Experimental Designs. *Practical Assessment, Research, and Evaluation*.
- He, Q. (2015). Computer-Assisted Language Learning and the Improvement on Reading Ability. *2015 International Conference on Logistics, Informatics and Service Sciences (LISS)*.
- Hu, Y., & Ojihara, M. (2011). Nextone Player: A Music Recommendation System Based on User Behavior. *Proceedings of the 12th International Society for Music Information Retrieval Conference*.
- Hubbard, P. (1996). Elements of CALL Methodology: Development, Evaluation, and Implementation. En Martha Pennington, *The Power of CALL* (págs. 15-32). Houston: Athelstan.
- Huertas, A., & Pantoja, A. (2016). Efectos de un programa educativo basado en el uso de las TICs sobre el rendimiento académico y la motivación del alumnado en la asignatura de tecnología de educación secundaria. *Educación XXI*, 19(2), 229-250.
- Huitt, W., Hummel, J., & Kaeck, D. (1999). *Internal and External Validity*. Obtenido de edpsycinteractive.org: <http://www.edpsycinteractive.org/topics/intro/valdgn.html>
- Instituto Mexicano para la Competitividad A. C. (2015). *Inglés es posible. Propuesta de una Agenda Nacional*. Recuperado el 12 de diciembre de 2018, de Instituto Mexicano para la Competitividad A. C.: [https://imco.org.mx/wp-content/uploads/2015/04/2015\\_Documento\\_completo\\_Ingles\\_es\\_posible.pdf](https://imco.org.mx/wp-content/uploads/2015/04/2015_Documento_completo_Ingles_es_posible.pdf)
- Isinkaye, F., Folajimi, Y., & Ojokoh, B. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 261-273.
- Leitner, S. (1972). So lernt man lernen: Der Weg zum Erfolg. *Herder*.

- Li, T., Jin, L., Wu, Z., & Chen, Y. (2019). Combined Recommendation Algorithm Based on Improved Similarity and Forgetting Curve. *MDPI Information Journal*, 10, 1-18.
- Lucas, S., & Meseguera, J. (2016). Normal forms and normal theories in Conditional rewriting. *Elsevier Journal of Logical and Algebraic Methods in Programming*, 67-97.
- Mishra, A., & Vishwakarma, S. (2015). Analysis of TF-IDF Model and its Variant for Document Retrieval. *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, 772-776.
- Moreno, D., Narvaez, C., Sastoque, S., & Garnica, G. (2016). Computational Model Based on Language Development Theories for Languages Learning and Training: Vocabulary Module. *2016 XLII Latin American Computing Conference (CLEI)*.
- Ozkan, S., & Koseler, R. (2009). Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*(53), 1285-1296.
- Papis, M., Kuczynski, P., & Seguin, W. (2020). *Ruby Version Manager (RVM)*. Obtenido de Ruby Version Manager (RVM): <http://rvm.io/>
- Peralta-Quintero, C.-A., Castro-Sánchez, N.-A., González-Serna, J.-G., & Nieto-Benítez, K. (2019). Un experimento para determinar el efecto de la curva del olvido en el aprendizaje del signo lingüístico. *3a Jornadas de Ciencia y Tecnología Aplicada*.
- Pham, X.-L., Chen, G.-D., Nguyen, T.-H., & Hwang, W.-Y. (2016). Card-based design combined with spaced repetition: A new interface for displaying learning elements and improving active recall. *Computers & Education*(98), 142-156.
- Ponnam, L., Deepak, P., Sreenivasa, N., Siva, N., & Yellamati, S. (2016). Movie recommender system using item based collaborative filtering technique. *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*.
- Puren, C. (2004). Del enfoque basado en tareas a la perspectiva de la cooperación. *Cahiers de l'APLIUT*, 23, 10-26.
- Real Academia Española. (2021). *Diccionario de La Lengua Española*. Recuperado el Septiembre de 2021, de Real Academia Española: <https://dle.rae.es>
- Restrepo Ramos, F. D. (2015). Incidental vocabulary Learning in second language acquisition: A literature review. *Profile: Issues in Teachers' Professional Development*, 157-166.
- Rice, P. (1997). *Desarrollo humano: estudio del ciclo vital*. Pearson Educación.
- Richardson, M., Burges, C., & Renshaw, E. (2013). *MCTest dataset and models*. Obtenido de GitHub: <https://github.com/mcobzarenco/mctest>
- Ridout, S. (2021). *ReadLang*. Recuperado el Noviembre de 2021, de ReadLang: <https://readlang.com/>
- RoadToGrammar. (2019). *Text Analyzer*. Obtenido de RoadToGrammar: <http://www.roadtogrammar.com/textanalysis/>
- Saussure, F. (2008). *Curso de lingüística general*. Buenos Aires: Losada.
- Schimanke, F., Ribbers, S., Mertens, R., & Vornberger, O. (2015). Implications of Short Term Memory Research for the Design of Spaced Repetition Based Mobile Learning Games. *2015 IEEE International Symposium on Multimedia (ISM)*.
- Schön, D. (1983). *The Reflective Practitioner: How Professionals Think in Action*.
- Secretaría de Educación Pública. (2004). *Manual de Estilos de Aprendizaje*. SEP, Subsecretaría de Educación Media Superior. Recuperado el 18 de abril de 2019, de Jalisco. Gobierno del Estado.: <http://edu.jalisco.gob.mx/cepse/cisneros-2004manual->

- de-estilos-de-aprendizaje-sep-subsecretaria-de-educaci%C3%B3n-m%C3%A9xico-superior
- Secretaría de Educación Pública. (2011). *Programa Nacional de Inglés en Educación Básica*. Recuperado el 12 de noviembre de 2018, de Gobierno de México: <https://www.gob.mx/aefcm/documentos/documentos-programa-nacional-de-ingles-en-educacion-basica> (último acceso 12 de noviembre del 2018)
- Secretaría de Educación Pública. (2017). *Estrategia Nacional de Inglés*. Recuperado el 12 de noviembre de 2018, de Gobierno de México: [https://www.gob.mx/cms/uploads/attachment/file/289658/Mexico\\_en\\_Ingle\\_s\\_DIGITAL.pdf](https://www.gob.mx/cms/uploads/attachment/file/289658/Mexico_en_Ingle_s_DIGITAL.pdf)
- Secretaría de Educación Pública. (2017). *Whiteboard. Estrategia Nacional de Inglés*. Recuperado el 12 de noviembre de 2018, de Gobierno de México: <https://www.gob.mx/sep/videos/estrategia-nacional-de-ingles>
- Secretaría de Educación Pública. (2021). *Subsecretaría de Educación Básica*. Recuperado el 7 de septiembre de 2021, de Gobierno de México: <https://educacionbasica.sep.gob.mx>
- SET Sociedad Española de Toxicomanías. (2006). *Tratado SET de trastornos adictivos*. Editorial Médica Panamericana.
- Swart, A. (2018). An epic consequence of classroom revision of course content for first-year African engineering students. *World Transactions on Engineering and Technology Education*, 16(4), 368-373.
- TALP Research Center. (2021). *Welcome | FreeLing Home Page*. Recuperado el 12 de diciembre de 2021, de Universitat Politècnica de Catalunya: <https://nlp.lsi.upc.edu/freeling/node/1>
- Test-English.com. (2021). *Test English A1 Reading Tests*. Recuperado el 12 de noviembre de 2021, de Test English: <https://test-english.com/reading/a1/>
- The PostgreSQL Global Development Group. (2020). *About PostgreSQL*. Obtenido de PostgreSQL: <https://www.postgresql.org/about/>
- Vargas, J. (2010). *Tabla de valores de probabilidad acumulada para la distribución normal estándar*. Recuperado el 12 de junio de 2021, de <https://jrvargas.files.wordpress.com/2010/07/tabla-z.pdf>
- Weblingua Ltd. (2021). *Text Inspector: Analyse the Difficulty Level of English Texts*. Recuperado el 12 de diciembre de 2021, de Text Inspector: <https://textinspector.com/>
- Wei, X., Limei, L., Liyuan, S., Xiaoli, Z., Lina, G., & Chunhe, L. (2018). E-Learning Pedagogical Strategies in traditional medical teaching. *9th International Conference on Information Technology in Medicine and Education*, 477-479.
- Wikipedia Foundation. (2021). *Simple English Wikipedia*. Recuperado el 12 de noviembre de 2021, de Simple English Wikipedia: [https://simple.wikipedia.org/wiki/Main\\_Page](https://simple.wikipedia.org/wiki/Main_Page)
- Wilde, O. (2001). *The Canterville Ghost and Other Stories*. Dover Publications.
- Wozniak, P. (2018). *SuperMemo 1.0 for DOS (1987)*. Recuperado el 12 de noviembre de 2021, de SuperMemo: [https://supermemo.guru/wiki/SuperMemo\\_1.0\\_for\\_DOS\\_\(1987\)#Algorithm\\_SM-2](https://supermemo.guru/wiki/SuperMemo_1.0_for_DOS_(1987)#Algorithm_SM-2)
- Yeh, M., Toshtar, A., Guertin, L., & Yan, T. (2016). Using Spaced Repetition and Gamification to Enhance K-12 Student Science Literacy With On-Demand Mobile Short Reads. *2016 IEEE Frontiers in Education Conference (FIE)*.
- Zangara, A., & Sanz, C. (2015). Importancia de las estrategias de autorregulación en el aprendizaje y sus derivaciones para la enseñanza. Análisis de un caso en Educación Superior Universitaria. *X Congreso de Tecnología en Educación & Educación en Tecnología*, 79-89.

- Zeng, L., & Lin, L. (2011). An Interactive Vocabulary Learning System Based on Word Frequency Lists and Ebbinghaus' Curve of Forgetting. *2011 Workshop on Digital Media and Digital Content Management*, 313-317.
- Zhang, Y., Jia, W., Zhu, C., & Song, Y. (2015). EVOV: A video recommendation system to support sustainable vocabulary learning. *2015 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*.
- Zheng, W. (2015). Design of Mobile Micro-English Vocabulary System Based on the Ebbinghaus. *2015 Eighth International Conference on Internet Computing for Science and Engineering (ICICSE)*.