





# Retención de vocabulario académico del italiano como lengua extranjera en una universidad pública de Colombia

## Retention of academic vocabulary in Italian as a foreign language at a public university in Colombia

Luis Enrique Pardo-Granados<sup>1</sup>   
Carlos Enrique Carrillo-Cruz<sup>2</sup>    
Santiago Ruíz-Bejarano<sup>3</sup> 

<sup>1</sup> Universidad Nacional, Bogotá, Colombia. [ipardog@unal.edu.co](mailto:ipardog@unal.edu.co)

<sup>2</sup> PhD. Universidad Libre. Universidad Nacional, Bogotá, Colombia. [carlosecarrilloc@unilibre.edu.co](mailto:carlosecarrilloc@unilibre.edu.co)

<sup>3</sup> Universidad Nacional, Bogotá, Colombia. [sruizb@unal.edu.co](mailto:sruizb@unal.edu.co)

Recibido: 17 de marzo de 2024

Aceptado: 14 de mayo de 2024

Publicado en línea: 12 de agosto de 2024

Editor: Matilde Bolaño García 

**Para citar este artículo:** Pardo Granados, L.E., Carrillo Cruz, C.E. y Ruíz Bejarano, S. (2024). Retención de vocabulario académico del italiano como lengua extranjera en una universidad pública de Colombia. *Praxis*, 20 (2), xx-xx.



## RESUMEN

La presente investigación mixta de diseño integrado busca determinar el porcentaje de vocabulario adquirido a corto plazo en clubes semanales de italiano II de una universidad pública de Colombia. Este artículo desarrolla, desde su marco metodológico, tres fases que comprenden el estudio de las necesidades, el componente cuantitativo y el cualitativo. El repertorio lingüístico seleccionado se centró en tres variables: frecuencia, complejidad y utilidad. Para la frecuencia se emplearon dos corpus del italiano escrito. Se realizaron encuestas sobre las posturas de los participantes frente a los clubes, *quizzes* y palabras, y los resultados se examinaron con la herramienta de análisis cualitativo Atlas.Ti para evidenciar las categorías cuantitativas. Las correlaciones entre las tres variables se determinaron con el programa estadístico SPSS. Los datos permitieron identificar la eficiencia en la adquisición de palabras mediante el uso de material auténtico a través de un corpus de italiano con fines académicos. De esta forma se confirma la importancia de enseñar vocabulario de forma consciente en los cursos de lengua y cómo la presentación de términos en múltiples contextos nativos permite su mejor comprensión.

**Palabras clave:** corpora; frecuencia; italiano; retención de vocabulario.

## ABSTRACT

This mixed research of integrated design seeks to determine the percentage of vocabulary acquired in the short term in weekly Italian II clubs at a public university in Colombia. This article develops three phases comprising the study of needs analysis, the quantitative component and the qualitative component. The selected linguistic repertoire focused on three variables: frequency, complexity and usefulness. Concerning frequency, two corpora of written Italian were used. Surveys were conducted on the participants' attitudes towards clubs, quizzes and words, and the results were examined with the qualitative analysis tool Atlas.Ti to highlight the quantitative categories. Correlations between the three variables were determined with the SPSS statistical program. The data made it possible to identify the efficiency of word acquisition by using authentic material through a corpus of Italian for academic purposes. This confirms the importance of consciously teaching vocabulary in language courses and how the presentation of terms in multiple native contexts allows for better comprehension.

**Keywords:** corpora; frequency; italian; vocabulary retention.



## INTRODUCCIÓN

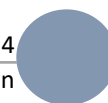
La particularidad con la que un individuo construye su realidad se manifiesta mediante la selección del repertorio lingüístico, cuya extensión impacta directamente tanto el nivel de complejidad como la descripción detallada de lo observable (Steiner y Russo, 2002). Por lo tanto, el vocabulario poseído resulta esencial ya que, a medida que se adquieren vocablos en diversas áreas, la representación del mundo será cada vez más sólida y se elaborarán pensamientos de una mayor complejidad; algo imposible de realizar sin la concreción del pensamiento crítico y sin la existencia y posterior estructuración coherente de palabras con miras a facilitar los procesos de enseñanza (Álvarez *et al.*, 2023). Asimismo, construir un repertorio lingüístico es la base fundamental para dominar otra lengua, ya que ningún conocimiento gramatical permite comunicarse efectivamente si no se dominan las palabras necesarias para expresar determinados conceptos (Rice y Tokowicz, 2020). Además, abordar otro idioma con un enfoque natural desligado de lo gramatical facilita la comprensión y el desarrollo del lenguaje (Kumaravadivelu, 2006).

A pesar de lo anterior, es evidente cómo se ha dejado de lado el enfoque lexical en la enseñanza. En primer lugar, como indica Milton (2009), dentro de la academia y el mundo de las lenguas extranjeras se concibe que un puñado de palabras son suficientes; no obstante, lo cierto es que se necesita un repertorio de miles de términos para desenvolverse tal como lo hace un hablante en su lengua materna (L1). Adicionalmente, la adquisición de vocabulario se ha percibido como un instrumento secundario por el enfoque dado a la enseñanza de L2, en el que se recae en aspectos gramaticales y estructurales en primera medida.

Así, esta investigación hace énfasis en la enseñanza del léxico académico. Cesati (2019), señala que un vocabulario académico puede ser abstracto, puede fungir para variadas funciones comunicativas en textos de registro formal e informal; que transmiten conocimientos sobre las más diversas áreas del saber, textos de divulgación o textos especializados, en los cuales el léxico académico acerca al lector a

los tecnicismos pertenecientes a las materias disciplinares. Más aún, las palabras académicas son útiles en la composición de textos extensos, formales y de diferente tipo que en los últimos años de secundaria se tienen que enseñar. Es de hecho, el léxico el que sirve para la realización de escritura escolar compleja, con la cual los alumnos desarrollan conocimientos y pensamientos en favor de una escritura que con justa razón se puede catalogar como académica, en tanto funciona desde registros sostenidos/elevados, precisos, usualmente despegados e impersonales. La enseñanza de una lengua debe también promover el aprendizaje de léxico académico, útil en las escuelas, universidades, prensa y más contextos formales, ya que para este enfoque existen categorías de corpus con mucho que contribuir a la enseñanza de lenguas extranjeras, directa o indirectamente (Huang, 2018). En este caso los de lengua escrita.

Por otro lado, existen estudios en plataformas de publicaciones indexadas como Scopus, ERIC Database, Academia o ResearchGate y libros divulgados por Oxford y Cambridge donde se corrobora una creciente rama investigativa preocupada por la enseñanza y retención de vocabulario, precisamente porque sin gramática es poco lo que puede ser expresado, pero sin vocabulario, nada puede serlo (Wilkins, 1972, p. 111). Dichos trabajos se centran en su gran mayoría en la lengua inglesa como objeto de estudio, pero han encontrado una pequeña contraparte en aquellos que abordan el italiano como lengua extranjera (ILE), como segunda lengua (ISL) e incluso como lengua materna (L1). El interés investigativo en la adquisición de léxico se debe al decreciente uso de vocablos técnicos y académicos por parte de las nuevas generaciones de universitarios y el impacto que esto tiene en la escritura y comprensión de textos. Una muestra de este fenómeno son las denuncias frente a la insuficiencia de una unidad didáctica fuera del contexto incidental que se presenta en el escenario académico de estudiantes de último año de bachillerato y universitarios (Cesati, 2019).



Sin embargo, enseñar desde un enfoque léxico no es algo que todavía resuena lo suficiente en la esfera didáctica. No es de extrañar que en Colombia no haya una gran cantidad de bibliografía en el tema. Tras realizar una búsqueda, no se encontró indagación sobre la enseñanza-adquisición de léxico italiano, ni frente a la orientación de un Corpus en la didáctica, en donde se realizó el estudio. Sin menoscabo de lo anterior, se puede decir que los Corpus están teniendo un considerable avance en Colombia, donde han tomado y siguen tomando fuerza con los siguientes enfoques: elaboración de un corpus computarizado de errores escritos con textos de estudiantes universitarios en Colombia por parte de la Universidad de Antioquia, y Cardicor: un corpus para uso de los traductores de medicina de la Universidad Nacional de Colombia. Se ha dado, por lo tanto, un uso distinto al diseño de métodos de enseñanza de una lengua con la colección de textos dispuestos en los corpus lingüísticos.

Actualmente, el uso de herramientas tecnológicas hace parte de las tendencias más importantes en la educación internacional (Mukhamadiarova et al., 2020). Los estudios demuestran que enfoques de aprendizaje basados en corpus encuentran mayores implicaciones para diseñar clases de vocabulario centradas en el estudiante de segunda lengua (Zhang, 2022). Sin embargo, hay ausencia de implementación, y por ende, de estudios de su uso como herramienta directa de enseñanza-aprendizaje por parte del profesorado en los colegios y en la educación superior; a sabiendas que múltiples autores reconocen la importancia y la creciente rama investigativa del corpus en la didáctica (Satake, 2022). Boulton y Cobb (2017) afirman que el *Data-Driven Learning* (DLL), esto es, el análisis de concordancias, palabras o unidades fraseológicas en diversas oraciones y el contraste de las formas en que el estudiante se comunica, puede incrementar la sensibilidad lingüística y la habilidad de trabajar con material auténtico. Este método ha propiciado logros en su mayoría más considerables en vocabulario y gramática con respecto otros (Vyatkina, 2020) y sus beneficios se han demostrado empíricamente en varios momentos, pero aun así

sigue siendo infravalorado por docentes de lenguas. Conrad (2000), discute el potencial que tiene el Corpus para revolucionar la enseñanza y Römer (como se citó en Vyatkina, 2020) diez años antes hizo notar la resistencia de los docentes a su implementación. La renuencia frente al DLL se debe en parte a la falta de guías y orientaciones para docentes y usuarios que quieran integrar esta herramienta en la enseñanza/aprendizaje dentro de plataformas que fueron diseñadas por lingüistas con interfaces muy técnicas y sin un objetivo pedagógico (Vyatkina, 2020). Aquí es donde la presente investigación se inclinó al uso de *hands-off activities*; donde los estudiantes trabajan con las concordancias directamente en el club sin involucrar su participación en la búsqueda y selección de estas, y no en *hands-on activities*; técnica instruccional en la que se aprende mediante la directa participación del aprendiz en la búsqueda y selección de material conduciendo sus propios experimentos, puesto que al trabajar en clubes de una hora por semana no se sacaría provecho de una previa instrucción hacia cómo navegar en el corpus; en cambio, se otorgaron diferentes fragmentos de texto con las palabras objeto en los quizzes.

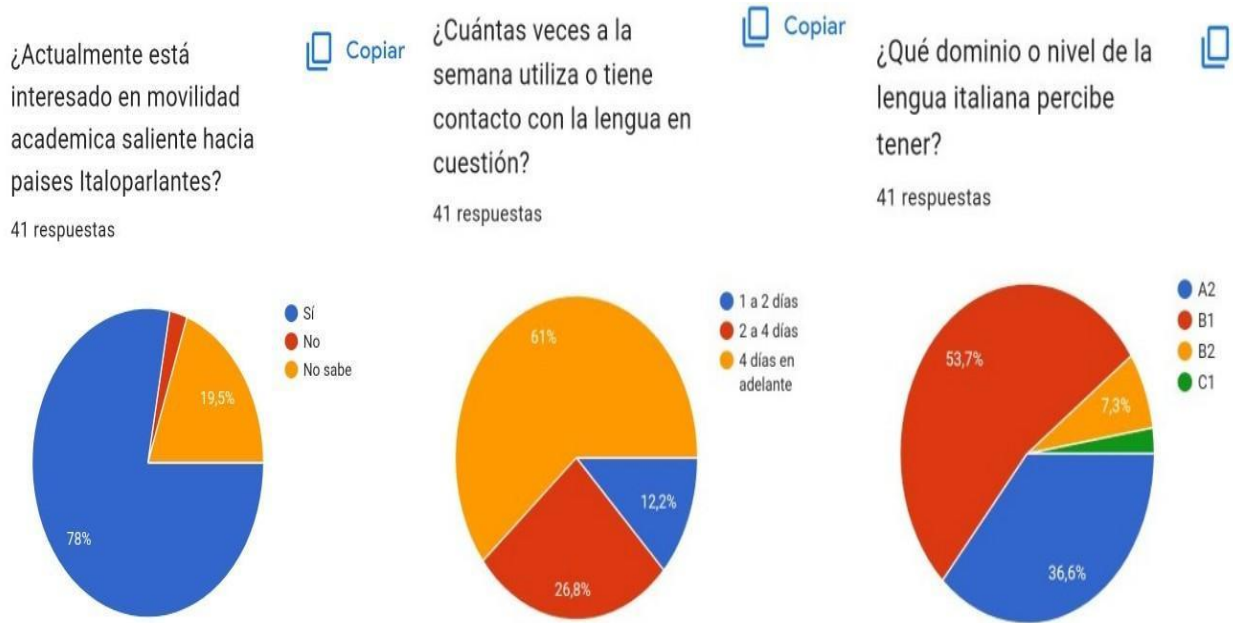
En el marco de la investigación se propuso medir la retención a corto plazo, en primera medida por la naturaleza de los clubes, de duración de una hora y una población generalmente cambiante. Luego, hay que tener en cuenta que las habilidades del aprendiente pueden ser específicamente representadas en distintas velocidades de aprendizaje y olvido, y cómo estas son tareas independientes que también deben ser recolectadas de forma independiente al momento de trazar un perfil del estudiante (Li, 2018). Y es que para garantizar una adquisición a largo plazo se tiene que pasar por el corto plazo de forma consciente e identificar fortalezas y debilidades.

Cuando los estudiantes culminan uno de los cursos de italiano de la universidad, comprenden la gramática; sin embargo, poseen un inventario de palabras restringido. Este fenómeno fue evidente a lo largo de la participación de estos sujetos en los clubes, donde si bien se podían tener cinco palabras

como objeto, quince terminaban por ser las aprendidas de forma incidental en las oraciones de ejemplo. La cantidad de términos adquiridos de esta manera varió por estudiante y siempre fue mayor a los vocablos objeto. Lo anterior muestra que los aprendices no podían expresar minuciosamente sus ideas en temas que en la L1 probablemente sí

puesto que “los límites de la lengua objeto son los límites del entendimiento de la realidad propia” (Wittgenstein *et al.*, 2001, p. 68). La figura 1 muestra los resultados de una encuesta acerca de los conocimientos previos de los miembros del grupo de WhatsApp de estudiantes de italiano de la universidad creado por el monitor.

Figura 1. Análisis preliminar de los participantes.



Fuente: Elaboración propia.

La encuesta evidencia un alto interés en movilidad saliente y la necesidad de reforzar el dominio de la lengua si los participantes han de hacer un intercambio en una universidad italiana, donde se les exija un registro alejado y técnico. Eventualmente, esta comunidad de alumnos, con bases de italiano y con interés en hacer intercambios académicos a países italo hablantes, debe conocer el registro formal para la lectura, la elaboración y la entrega de trabajos académicos de manera que puedan cumplir con las exigencias de la institución con la que se tiene convenio, o para adentrarse en temas que denotan y requieren un más alto grado de erudición en la lengua objeto.

El nuevo vocablo existe cual semilla fresca lanzada en el terreno de la discusión (Wittgenstein, L, 1980). Así, en este marco contextual, surgió el interés de llevar a cabo un estudio donde se promoviera el

léxico italiano en una comunidad universitaria hispanohablante con miras a dar solución a problemáticas específicas en el campo de la educación (Carrillo, 2020).

La presente investigación mixta pretende entonces medir la retención del vocabulario académico por parte de los participantes de los clubes de Italiano II. A continuación, se discuten la metodología de las actividades y de la recolección de datos y su posterior evaluación. Se describe, por una parte, el análisis cuantitativo en Excel y, por otra, el análisis cualitativo mediante Atlas.Ti para el contraste de los resultados con los cuestionarios de percepción de los participantes. Finalmente, se reconoce si hay o no correlación entre las tres variables aplicadas al repertorio lexical seleccionado (frecuencia, complejidad y utilidad) con el uso del programa estadístico multivariable SPSS.

## El repertorio lingüístico

Los clubes iniciaron con una lectura de las varias definiciones de una palabra y sus ejemplos en oraciones; fuentes oficiales como diccionarios (Treccani, Nuovo Mauro); páginas de oraciones de ejemplos (frazodict, dazy) fueron usadas. Luego se pasaba a una actividad de escucha con un audio de 1 a 2 minutos que ponía en contexto las palabras vistas (usualmente 5), acompañada de una ayuda visual ya que (...) el desempeño de los oyentes en los exámenes de comprensión auditiva apoyada con imágenes estáticas es mejor que la de aquellos con el mismo audio apoyados de guiones parafraseados (Ginther, 2002; Ockey, 2007, como se citó en Maleki y Safaee, 2011, p. 8), para resolver preguntas en grupo. Finalmente, se daba lugar al quiz en los computadores, que a través de distintos tipos de preguntas hizo uso de fragmentos de texto del Corpus di Bologna dell'Italiano Scritto (CODIS) con las palabras objeto.

Con base en tres variables (frecuencia, complejidad y utilidad), se seleccionaron ciertos términos para determinar la relación de los estudiantes con ellos. Estos parámetros atienden al modelo de He y Godfroid (2018), quienes proponen un nuevo método para la creación de listas de vocabulario por enseñar.

La primera variable, la frecuencia de palabras, se determinó a través del CODIS, donde la unidad de medición es ocurrencias por millón o el número de veces que aparece la palabra dentro de un corpus de 165 millones de palabras y el Lessico dell'Italiano Scritto (LIS) con 25M de ocurrencias. Hay que tener en cuenta que tanto nativos como no nativos tienen una lista preconfigurada de vocablos que tienden a aprender en una primera etapa (Milton, 2009), y es ahí que muchas de las aplicaciones de corpus lingüístico, y libros para la enseñanza de lengua extranjera se enfocan en los de alta incidencia. Una frecuencia menor tiende a relacionarse con una adquisición más lenta. Sin embargo, esto no las hace prescindibles y es posible, si necesario, incrementar en el aula su aparición a través del aprendizaje distribuido; que consiste en localizar el contenido en diversas posiciones no centralizadas, y así permitir

un aprendizaje íntegro (Kasahara y Kanayama, 2011) de palabras consideradas como útiles. Así, a medida que el nivel del estudiante mejora, se llega a un balance donde las palabras menos frecuentes pasan a ser bien conocidas (Milton, 2009).

En efecto, la actividad de audio permitió retomar palabras vistas en sesiones anteriores junto con las del club correspondiente, de manera que el conocimiento previo continuó ajustándose y permitió la construcción de nuevas estructuras (Piaget, 1970). En todo caso, lo que más importa en este aprendizaje es que la selección sea contextualizada y “didactizada” en función de las necesidades lingüísticas de los alumnos (Cesati, 2019). La segunda variable, la complejidad, alude a una serie de factores lingüísticos de cada lengua que determinan el nivel de dificultad de aprendizaje. Al momento de aprender listas de vocablos, una mayor imaginabilidad de la palabra (el grado con el cual genera una imagen mental), hace que sea más posible recordarla, de esto sigue que la concreción de la palabra afecta su adquisición. También la longitud de esta medida en cantidad de sílabas, la cercanía de la estructura morfológica y deletreo con respecto a la L1 (Ellis & Beaton, 1993), y si el deletreo de esas palabras es frecuente en la L2 o inusual, algo que puede incluso desacelerar la adquisición por parte de hablantes nativos. En el caso concreto del italiano, la semejanza escrita del repertorio léxico con vocablos del español es evidente, lo que dificulta una correcta producción textual. Por lo tanto, las preguntas 3 y 4 del *quiz* contrastaron palabras y pseudopalabras similares ortográficamente, considerando que de este modo las probabilidades de correcta adquisición aumentan (Baxter *et al.*, 2022). Asimismo, entre los métodos empleados en la presentación de palabras para comprender y recordar está la concientización de su origen literal, técnica llamada elaboración etimológica (Meara, 2009). Además, se trabajó con la morfología, que examina cómo se forman las palabras con prefijos, sufijos y lexemas y cumple un rol importante en el aprendizaje de idiomas (Zhang, 2022). Esto lleva a pensar en el lema modificado, que también cuestiona la baja frecuencia como variable absoluta de dificultad,



pues palabras como “evitar” permitirían rápidamente aprender “evitable” e “inevitable”.

Laufer y Nation (2012), por su parte, habla de una tercera variable: utilidad, y menciona que existen dos tipos de palabras importantes a seleccionar. En general, se dice que una palabra es útil cuando es frecuente en una lengua y por ende se puede usar para cualquier propósito. Sin embargo, las que son de baja frecuencia también son útiles en la medida que su uso recae en las necesidades particulares de los hablantes, pues existe un campo semántico específico del área de especialidad de cada individuo que es fundamental para el campo laboral y académico. En esa medida, repetir varias veces el proceso de enfrentarse a estos términos concretos incrementa las probabilidades de procesarlos (Schuetze, 2017). Como afirma Paul Nation (2016), uno de los criterios en las listas de vocabulario hace referencia a que los hablantes nativos consideran muy útiles, palabras que no siempre tienen alta frecuencia. Aunque algunas palabras especializadas (técnicas) aparezcan esporádicamente en los corpus de lengua o en el día a día, pueden ser altamente frecuentes en un área y son verdaderamente útiles para la comunicación dentro de ese campo.

Para la utilidad y complejidad de las 39 palabras escogidas por el monitor (nativo), se optó por una escala de medición del 1 a 7, con base a estudios realizados por Tidball y Treffers-Daller (2008) y Simpson-Vlach y Ellis (2010), en la que el valor 1 corresponde a poco útil/fácil, y 7, a muy útil/difícil. Estos términos fueron evaluados por docentes de italiano de la universidad, pues el apoyo de expertos se ha demostrado eficiente en conjunto con el uso de corpus por la capacidad flexible y versátil de la intuición humana, lo que exige su intervención (He y Godfroid, 2018).

## METODOLOGÍA

El objetivo general del estudio es medir el porcentaje de vocabulario adquirido por parte de los estudiantes que asisten a los clubes de italiano-intermedio. Los objetivos específicos derivados son los siguientes:

- Mostrar la relación entre las 3 variables empleadas en la selección de las palabras objeto y cómo influyen en el aprendizaje de nuevas palabras.
- Reconocer qué factores influyen en la adquisición de vocabulario basado en el testimonio de la población del estudio.
- Fomentar el uso de Corpus para el aprendizaje y adquisición de nuevas palabras que se puedan emplear en contextos académicos y literarios.

En relación con el método, se optó por abordar un diseño mixto integrado, en el cual la información cuantitativa y la cualitativa se recopilan de forma simultánea o secuencial, de modo que un tipo de información cumple una función complementaria y sustenta los hallazgos de la anterior (Creswell, 2011). En este caso, se empezó recolectando datos cualitativos concernientes al conocimiento *a priori* de los participantes, es decir, su grado de inmersión en la lengua, y luego se dio paso a la realización de los clubes y los *quizzes* de vocabulario. Al finalizar el estudio, se llevó a cabo una encuesta sobre las percepciones y el valor que la población de estudio asignaría al proceso de adquisición de vocabulario, además del posterior análisis de los resultados.

Así las cosas, el desarrollo de la investigación se concretó principalmente en tres fases:

1. Una vez establecida la problemática, se determinaron los instrumentos y la población. En seguida, tanto docentes como estudiantes realizaron la encuesta sobre inmersión en lengua italiana.
2. La implementación incluyó actividades en clase, donde se enseñaron las palabras en contexto, y posteriormente la realización de los *quizzes*.
3. El análisis de los datos tras la implementación de los *quizzes* de vocabulario y de las respuestas obtenidas en los formularios de percepción.

La población objeto del estudio comprendió un grupo de estudiantes adultos pertenecientes a la comunidad universitaria de diferentes edades con un nivel de italiano pre-intermedio en su mayoría, quienes mostraban interés en mejorar su desempeño con dicha lengua.

La población objeto del estudio comprendió un grupo de estudiantes adultos de diferentes edades pertenecientes a la comunidad universitaria con un nivel de italiano pre-intermedio en su mayoría, quienes mostraban interés en mejorar su desempeño con dicha lengua.

El estudio de campo se realizó dentro de las instalaciones de la universidad, en el aula, donde se llevaron a cabo varios procedimientos dirigidos a facilitar el aprendizaje del léxico. El análisis preliminar de la percepción estudiantil sobre los conocimientos y las expectativas con la lengua se elaboró antes de los clubes por medio de un formulario. Por otra parte, se contó con la colaboración de cuatro docentes de italiano disponibles para la evaluación de las palabras en cuanto a su utilidad y complejidad.

La muestra estuvo compuesta por 45 participantes que accedieron a hacer parte del estudio durante la primera fase, en la cual se recolectaron un total de 45 respuestas bajo el consentimiento informado. También se recogieron 63 respuestas a través de los *quizzes* de retención de vocabulario a lo largo de siete clubes, considerando que la persona que participaba en más de un club no tenía que volver a diligenciar el consentimiento informado. Asimismo, se llevaron a cabo dos *quizzes* en clases distintas de italiano II para diversificar la población.

De igual modo, se obtuvieron testimonios a partir del grupo de personas que participaron en el proceso de los clubes más de un 50 %, así como de diez estudiantes de una de las clases de italiano II anteriormente nombradas. Los principales instrumentos para la recolección de datos fueron los siguientes:

- Corpus Italiano: Colección de diversos textos encontrados en el CODIS, de los cuales se extraen las palabras objeto y algunos fragmentos de texto. Asimismo, se usó para determinar la frecuencia por palabra junto con el Corpus LIS.
- Cuestionario de conocimientos preliminares a docentes y estudiantes: Desempeño, experiencia e inmersión en la lengua italiana. Además de indicar si posee alguna certificación. En el caso de los estudiantes se indaga sobre su contacto con la lengua, el nivel que perciben poseer y sus intereses académicos con el idioma.
- Formulario Utilidad y Complejidad: asignado a los docentes que accedieron a participar, en el cual se hace uso del criterio de los expertos al presentar la lista de palabras usadas por club y se pide asignar un puntaje del 1 a 7 (puede haber decimales) para cada variable en un total de 7 formularios por docente.
- Quizzes de retención de vocabulario: Siete preguntas que miden la retención de vocabulario a corto plazo con una puntuación máxima de 10 puntos, las cuales son de distinta tipología: selección múltiple, filling the gap, preguntas de inferencia, construcción de una oración, etc. Cabe resaltar, que el uso de corpus, especialmente el uso de concordancias contribuye a completar preguntas de tipo llenar el vacío, mucho más que las actividades tradicionales (Boulton, 2009).
- Cuestionario Final: Indagación sobre las percepciones de los participantes respecto a lo que aprendieron y cómo.

La utilización de varias herramientas buscaba el contraste entre los datos cuantitativos y los cualitativos: los resultados, el criterio docente y la percepción estudiantil. En esa medida fue relevante la implementación de cuestionarios/formularios a lo largo de la investigación. Así, por ejemplo, los *quizzes* de retención de vocabulario integraron





preguntas de distinta tipología teniendo en cuenta que, según Nurmukhamedov (2012), un cuestionario de alta calidad conlleva un alto grado de redacción y contenido. Esto último es importante puesto que este tipo de preguntas son necesarias para lograr una transición desde un aprendizaje guiado hasta el acto comunicativo en sí mismo. Como indica Ur (2012), en el momento de aprender la gramática de un idioma, las personas deben pasar por un proceso donde las respuestas sean controladas, y eventualmente, a medida que se familiarizan con la lengua, se les permite crear su propio discurso sin la necesidad de alguna instrucción.

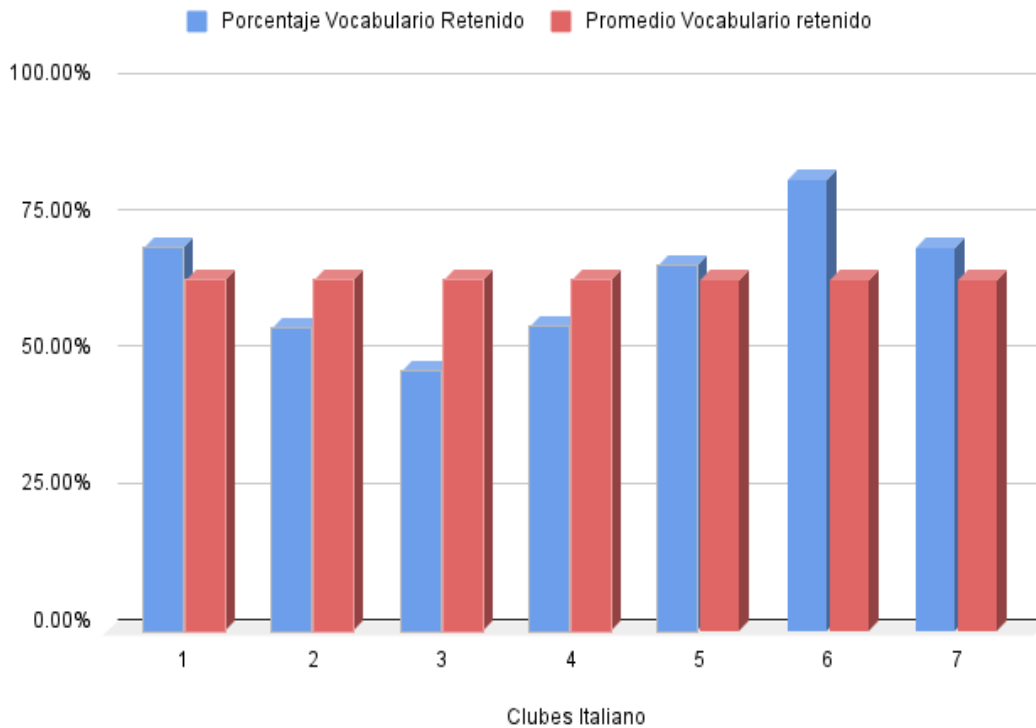
## RESULTADOS

Los resultados del estudio se muestran con respecto a las variables definidas frente a la retención de vocabulario, las cuales se midieron con base en las respuestas a los *quizzes* en cada club a lo largo de siete semanas. De esta forma se obtuvo una media de retención del 63,94 %. En concreto, los clubes

tres y seis registraron la menor y la mayor adquisición de vocabulario, 47,50 % y 82,50 % respectivamente.

Tomando en cuenta lo anterior, hay que considerar que los mejores resultados se mostraron en el club virtual realizado en sesiones remotas para la interacción con los participantes, el cual no dispuso de un protocolo de seguridad que detectara el uso de herramientas facilitadoras para la traducción. Sin embargo, a pesar de no tener suficiente evidencia para determinar si la virtualidad es mejor o no para este tipo de procedimientos pedagógicos que aplican el uso de herramientas tecnológicas, es un hecho que los crecientes beneficios de la inteligencia artificial (IA) en el campo de la educación deben ser utilizados, explotados y potenciados con fines específicos que contribuyan al desarrollo de habilidades. De igual forma, la virtualidad es una modalidad flexible que evidenció una mayor participación y disposición en el momento de realizar las actividades.

Figura 2. Porcentaje Retención de Vocabulario por Clubes.



Fuente: Elaboración propia.



Para determinar y contrastar el valor de las 3 variables con las que se miden las palabras objeto se consideró el uso del programa estadístico multivariable SPSS. Se optó por la herramienta de bootstrap analysis con un grado de confidencialidad

del 95%, cuya principal ventaja es reducir el sesgo analítico considerando el tamaño limitado de la muestra y la posibilidad de usar esta herramienta para realizar muestreos aleatorios.

**Tabla 1.** Estadística descriptiva bootstrap de las 39 palabras objeto.

Estadística descriptiva						
Estadística			Sesgo	Desviación Típica (TP)	BCa 95 % Intervalo de confianza	
					Mínimo	Máximo
Frecuencia	N	39	0	0		
	Mínima	9,00				
	Máxima	681,00				
	Media	165,8718	0,9369	21,4386	121,2731	214,5122
	Desviación Estándar	133,09656	3,70007	24,46577	93,07770	170,59768
Utilidad	N	39	0	0		
	Mínima	3,00				
	Máxima	6,00				
	Media	4,4944	-0,0003	0,1349	4,2353	4,7533
	Desviación Estándar	0,86314	-0,01582	0,06802	0,73708	0,95404
Complejidad	N	39	0	0		
	Mínima	3,00				
	Máxima	6,43				
	Media	4,7449	0,0025	0,1260	4,4945	5,0023
	Desviación Estándar	0,78576	-0,01498	0,08525	0,61869	0,91294
N válido (por lista)	N	39	0	0		

**Nota:** Elaboración propia que denota la relación de palabras que se realiza respecto a 1.000 desde Bootstrap.

**Fuente:** Elaboración propia.

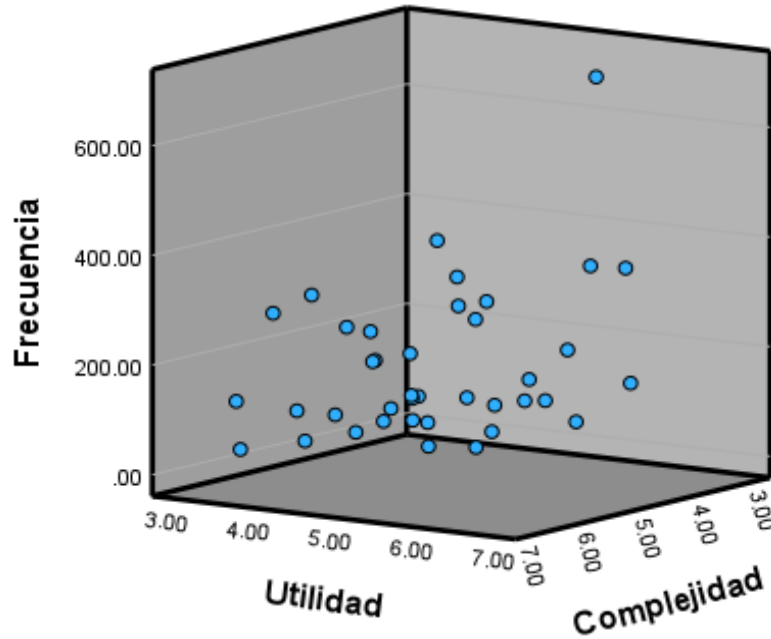
Por consiguiente, la frecuencia promedio de las 39 palabras analizadas fue de 165,8718 opm (DE=133,09656), asimismo utilidad y complejidad fueron 4,4944 (DE: 0,86) y 4,7449 (DE: 0,78) correspondientemente, en una escala de 1 a 7 puntos (valores asignados por los docentes). Los

docentes primero calificaron la utilidad de todas las palabras, 1 para fácil y 7 para difícil. Los resultados muestran una alta frecuencia tomando en cuenta que se trata de vocabulario académico y por ende poco usual dentro del repertorio lexical de los participantes, además tomando en cuenta el

tamaño del corpus empleado: CODIS (165 Mw). Por otra parte, se puede observar que las otras variables presentan valores moderados. Aunque haya palabras que demandan una mayor atención, un

gran número se encuentra dentro de un grado de dificultad aceptable tomando en cuenta el nivel de italiano de la población en el momento de realizar el estudio.

Figura 3. Modelo en 3D scatterplot para las 39 palabras del estudio.



Fuente: elaboración propia.

Por último, se analizaron las palabras como unidades individuales; es decir, no hubo casos de colocaciones, verbos frasales, perífrasis o cualquier unidad fraseológica. Como se observa en la figura 3, los términos poco frecuentes tienden a una menor utilidad y mayor complejidad. Sin embargo, aquellos vocablos de alta ocurrencia por sí solos no siempre corresponden a un mayor grado de utilidad funcional (Simpson-Vlach y Ellis, 2010), como se verá en el apartado cualitativo.

Algunas palabras, aunque hayan sido de poca frecuencia, fueron mejor asimiladas por la población dado que uno de los factores que determinan el paso de aprendizaje de L1 a L2 son los cognados, pues se asemejan al léxico ya existente dentro del repertorio del estudiante (Bardel *et al.*, 2012). De este modo, una palabra infrecuente en un corpus puede ser transparente en la lengua materna o un préstamo lingüístico conocido (extranjerismos, en el

caso de que los idiomas no estén relacionados), lo que hace que dicho término sea fácil de adquirir, a la vez que sus derivaciones, si bien pueden ser de baja frecuencia, resultan también asimilables (Laufer, 2014).

### Cualitativo

Se hizo uso de google forms y nearpod para recolectar los datos cualitativos de las categorías utilidad y complejidad. La variable de frecuencia está por fuera de este contraste ya que es una categoría que se tomó con los 2 corpora y no con el criterio docente, aunque bien es cierto que las tres variables se interrelacionan de una u otra forma.

Se tomaron en cuenta 3 tipologías de documentos de percepción a analizar:

1. Percepción de docente de italiano II.

2. Percepción de estudiantes de la clase de italiano II que participaron del Quiz 2 y
3. Percepción de los estudiantes más recurrentes en los clubes de italiano de Explora.

Explorar la percepción del docente de italiano II permite acercarse a la manifestación de actividad en términos de complejidad y utilidad del léxico explorado. A su vez, adentrarse en la interpretación estudiantil puede ser focalizada en momentos sobresalientes por analizar. Tanto el club dos como el tres fueron elaborados en espacios de clase brindados por docentes de italiano de la universidad. El club tres fue el de menor retención por la complejidad y frecuencia de algunas de sus palabras. Para este caso se contactó al docente de la asignatura italiano II del Departamento de Lenguas Extranjeras, quien incluso dio continuidad a la temática de la actividad con audio en la subsecuente clase por el interés de los participantes y la dificultad de la temática. La grabación utilizada fue de registro poético, haciendo alusión a la obra de Alessandro Manzoni: *I promessi sposi*. Por consiguiente, se volvió a trabajar con las palabras objeto, esta vez con el audio y su transcripción.

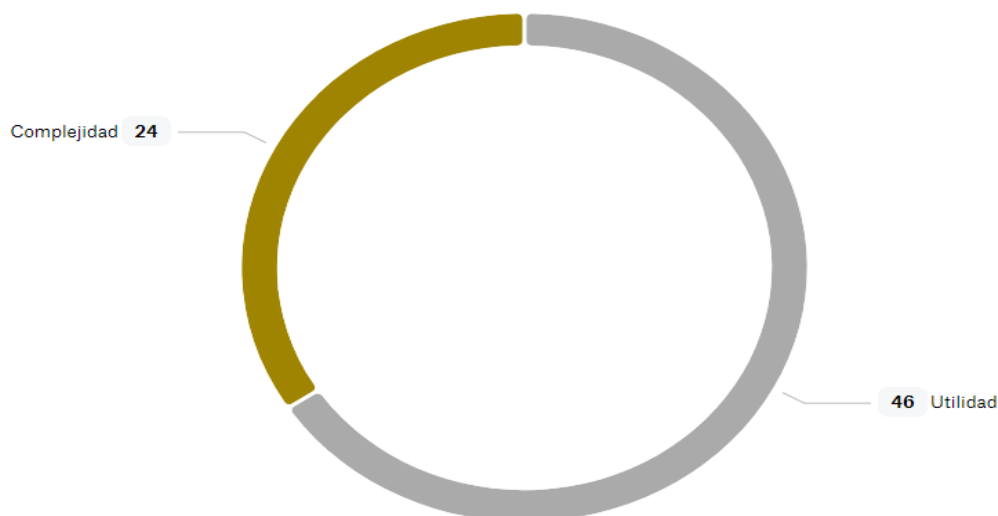
Algo dicho anteriormente es que encontrarse varias veces con el léxico es algo necesario para una

comprensión y adquisición real; máxime si ciertos términos se perciben difíciles, como mostraron los resultados del *quiz* de esta actividad, que fue el de mayor dificultad y menor retención. De ahí que fue preciso contrastar la percepción de los estudiantes con los resultados y ver su posición frente a las dos categorías mencionadas.

Finalmente, emergieron también percepciones considerablemente extensas de los participantes más recurrentes a los clubes, cuyo rendimiento en los *quizzes* se mantuvo en un nivel más alto que la media de la población estudiada, que, como se observó, consideró estar entre un A2 y un B1 mayoritariamente.

Para determinar la relación porcentual entre oraciones y categorías, se hizo uso de la herramienta de análisis cualitativo Atlas.Ti. De esta forma se encontró un total de 56 registros (oraciones) en los tres documentos analizados: 8 en el documento de percepción docente, 12 en el de percepción de estudiantes del club dos, y 36 en el de estudiantes más recurrentes a los clubes. Dentro de estos registros se hacen 24 menciones a la complejidad, lo que corresponde a un porcentaje de  $C=34,2857143\%$ , y 46 a la utilidad, con  $U=65,7142857\%$  de representación (figura 4).

Figura 4. Número de categorías en los documentos recolectados.



Fuente: elaboración propia.



Se encuentra homogeneidad en las percepciones sobre las dos categorías analizadas en el apartado cualitativo, y se reconoce una considerable complejidad en las palabras vistas por razones como el poco uso, abstracción, longitud y confusión con palabras ortográficamente similares tanto en la L1 como en la L2. Sin menoscabo de lo anterior, se confirma la existencia de una alta utilidad para el léxico en determinados contextos como el académico y en el de lengua escrita, lo cual generó motivación a la hora de aprender, siendo los participantes de naturaleza universitaria. Por lo tanto, tomar en consideración la enseñanza intencional y consciente del léxico surge como un deber de la educación lingüística para cada vez más docentes de italiano (Cesati, 2019).

## CONCLUSIÓN

La presente investigación propuso determinar el porcentaje de vocabulario adquirido por los participantes de los clubes de italiano II. Tras el análisis de los datos, se halló una variación de léxico para cada club con respecto a su retención. En consecuencia, los resultados de los usuarios dependen principalmente de la correlación entre las tres variables y por tanto fluctúan de acuerdo con la selección misma de palabras.

El promedio de retención a corto plazo en todos los clubes fue del 63,94 %, lo cual podría indicar que un determinado número de palabras por reforzar por club de conversación puede promover un nivel de aceptabilidad superior a la mitad de la terminología empleada en las sesiones. De esta forma, un docente podría hacer una proyección más reducida del vocabulario contraponiendo el avance curricular con el respectivo número de palabras aprendidas en promedio. Por lo tanto, el uso exclusivo de la variable de frecuencia para seleccionar qué léxico enseñar puede ser equívoco, máxime a la hora de transmitir el vocabulario de un área de trabajo específica.

Asimismo, para que los corpus de lengua mejoren su fiabilidad, deben seguirse nutriendo tanto de textos antiguos como modernos y de una posterior implementación de nuevas variables que trabajen a

tiempo con la frecuencia. También se puede afirmar que el léxico de registro elevado tiene una menor frecuencia, pero es necesario para expresarse sobre temas complejos y así entrar a ser parte de una comunidad académica. En ese sentido, puede ser positivo y posible acelerar conscientemente la adquisición de listas de vocabulario de un área de interés para enseñar con contexto o sin este, de forma directa, y no solo recibir palabras de forma incidental-natural con base en sus verdaderas tendencias de frecuencia, lo cual tomaría más años.

Así pues, se reafirma la necesidad de una ampliación guiada del vocabulario, constante y bien programada a través de todo el sendero educativo, como ya lo esperan las actuales indicaciones nacionales y las cartillas del Ministero dell'Istruzione (Cesati, 2019). Sin embargo, este proceso solo puede adelantarse una vez que el aprendiz haya logrado comunicarse y entender la gramática de la L2. De igual modo, se evidencia homogeneidad en las percepciones sobre las dos categorías analizadas en el apartado cualitativo por cuanto las palabras enseñadas tienen una considerable complejidad de acuerdo con las razones de frecuencia y de abstracción.

Este artículo da cuenta del primer estudio realizado en una universidad pública de Bogotá, Colombia, que aborda la enseñanza/retención del léxico italiano con enfoque basado en corpus. Considerando la creciente relevancia que recibe el aprendizaje de vocabulario dentro de la comunidad académica, en especial tomando en cuenta que el Ministero dell'Istruzione de Italia busca incentivar la adquisición y el empleo de nuevas palabras en contextos académicos, la presente investigación busca orientar la elaboración de clubes para el aprendizaje desde un enfoque léxico y guiar a docentes y administrativos en la elaboración de listas de vocabulario. Se espera que la audiencia tome como referencia los métodos, la metodología y los datos proporcionados para aplicarlos dentro del aprendizaje personal y la enseñanza de un idioma extranjero.





## DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores declaran que no tuvieron ningún interés personal en la revisión sistemática y el desarrollo de este artículo y que no se integraron conductas y valores inapropiados que diferían éticamente de los utilizados en el estudio. Por lo tanto, declaran que no existe ningún conflicto de intereses.

## REFERENCIAS BIBLIOGRÁFICAS

Álvarez, E., Collado, J. y Orefice, C. (2023). Pensamiento complejo para mejorar el perfil profesional docente en Colombia. Un estudio piloto de investigación transdisciplinar dentro de una red interuniversitaria. *Documentación de las Ciencias de la Información*, 2(46), 161-168.  
<https://doi.org/10.5209/dcin.87916>

Bardel, C., Gudmundson, A. y Lindqvist, C. (2012). Aspects of Lexical Sophistication in Advanced Learners' Oral Production: Vocabulary Acquisition and Use in L2 French and Italian. *Studies in Second Language Acquisition*, 34(2), 269-290.  
<https://doi.org/10.1017/S0272263112000058>

Baxter, P., Bekkering, H., Dijkstra, T., Droop, M., Van den Hurk, M. y Leoné, F. (2022). Contrasting orthographically similar words facilitates adult second language vocabulary learning. *Learning and Instruction*, 80, 101582.  
<https://doi.org/10.1016/J.LEARNINSTRUC.2022.101582>

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37-54.  
<https://doi.org/10.1017/S0958344009000068>

Boulton, A. y Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 67(2), 348-393.  
<https://doi.org/10.1111/LANG.12224>

Carrillo, C. E. (2020). El filtro afectivo positivo como estrategia de estimulación neurocognitiva para la modificación de conductas negativas en un niño en condición de discapacidad. *Universidad Libre*.  
<https://doi.org/10.18041/978-958-5578-44-9>

Cesati, F. (2019, del 12 al 14 de octubre). *Scrivere nella scuola oggi: obiettivi, metodi, esperienze: atti del 2. Convegno nazionale ASLI Scuola*, Siena, Università per Stranieri.

Conrad, S. (2000). Will Corpus Linguistics Revolutionize Grammar Teaching in the 21st Century? *TESOL Quarterly*, 34(3), 548-560.  
<https://doi.org/10.2307/3587743>

Creswell, J. W. (2011). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*.

Ellis, N. C. y Beaton, A. (1993). Psycholinguistic Determinants of Foreign Language Vocabulary Learning. *Language Learning*, 43, 559-617.

He, X. y Godfroid, A. (2018). Choosing Words to Teach: A Novel Method for Vocabulary Selection and Its Practical Application. *TESOL Quarterly*, 53(2), 348-371. <https://doi.org/10.1002/TESQ.483>

Huang, L. S. (2018). Taking Stock of Corpus-Based Instruction in Teaching English as an-International Language. *RELC Journal*, 49(3), 381-401.  
<https://doi.org/10.1177/0033688217698294>

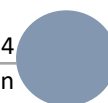
Kasahara, K. y Kanayama, K. (2021). When to conduct a vocabulary quiz, before the review or after the review? *System*, 103, 102641.  
<https://doi.org/10.1016/J.SYSTEM.2021.102641>

Kumaravadivelu, B. (2006). *Understanding Language Teaching: From Method to Postmethod*. Routledge.

Laufer, B., & Nation, I. S. P. (2012). Vocabulary. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163-176). London, England: Routledge.  
<https://doi.org/10.4324/9780203808184.ch10>

Laufer, B. (2014). Vocabulary in a second language: Selection, acquisition, and testing: A commentary on four studies for JALT vocabulary SIG. *Vocabulary Learning and Instruction*, 3(2), 38-46.  
<https://doi.org/10.7820/vli.v03.2.laufer>

Li, L. (2018). Sentiment-enhanced learning model for online language learning system. *Electronic Commerce Research*, 18(1), 23-64.  
<https://doi.org/10.1007/S10660-017-9284-5>



Maleki, A. y Safaee, M. (2011). The effect of visual and textual accompaniments to verbal stimuli on the listening comprehension test performance of Iranian high and low proficient EFL learners. *Theory and Practice in Language Studies*, 1(1), 28-36.

Meara, P. M. (2009). *Connected Words: Word associations and second language vocabulary acquisition*.

Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. *Multilingual Matters*.

Mukhamadiarova, A. F., Caserta, L. F., Kulkova, M. A. y Reuter, K. (2020). Use of Corpus Technologies for the Development of Lexical Skills. *Utopía y Praxis Latinoamericana*, 25(Extra-7), 185-194. <https://doi.org/10.5281/ZENODO.4009657>

Nation, I. S. P. (2016). *Making and Using Word Lists for Language Learning and Testing*. <https://doi.org/10.1075/Z.208>

Nurmukhamedov, U. (2012). Insights into non-native vocabulary teaching and learning. Rubén Chacón-Beltrán, Cristián Abello-Contesse, and Torreblanca-López María del Mar (Eds.). Bristol, UK: *Multilingual Matters*, 2010. Pp. v + 230. *Studies in Second Language Acquisition*, 34(1), 166-167. <https://doi.org/10.1017/S0272263111000593>

Piaget, J. (1970). Piaget's theory. En P. H. Mussen (Comp.), *Carmichael's manual of child psychology* (Vol. 2). Wiley.

Rice, C. A. y Tokowicz, N. (2020). A review of laboratory studies of adult second language vocabulary training. *Studies in Second Language Acquisition*, 42(2), 439-470. <https://doi.org/10.1017/S0272263119000500>

Satake, Y. (2022). The effects of corpus use on L2 collocation learning. *JALT CALL Journal*, 18(1), 34-53. <https://doi.org/10.29140/JALTCALL.V18N1.520>

Schuetze, U. (2017). Efficiency in Second Language Vocabulary Learning. *Die Unterrichtspraxis/Teaching German*, 50(1), 22-31. <https://doi.org/10.1111/TGER.12019>

Simpson-Vlach, R. y Ellis, N. C. (2010). An Academic Formulas List: New Methods in Phraseology Research. *Applied Linguistics*, 31(4), 487-512. <https://doi.org/10.1093/APPLIN/AMP058>

Steiner, G. y Russo, E. (2002). *Extraterritorial: ensayos sobre literatura y la revolución lingüística*.

Tidball, F. y Treffers-Daller, J. (2008). Analysing lexical richness in French learner language: what frequency lists and teacher judgements can tell us about basic and advanced words. *Journal of French Language Studies*, 18(3), 299-313. <https://doi.org/10.1017/S0959269508003463>

Ur, P. (2012). *A Course in English Language Teaching* (2.a ed.). Cambridge University Press. <https://doi.org/10.1017/9781009024518>

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17, 65-83.

Vyatkina, N. (2020). Corpora as open educational resources for language teaching. *Foreign Language Annals*, 53(2), 359-370. <https://doi.org/10.1111/FLAN.12464>

Wilkins, D. A. (1974). *Linguistics in Language Teaching* (1ª ed. australiana). Edward Arnold (Australia). (Publicado originalmente en 1972 por Edward Arnold, Londres). ISBN: 0726720070

Wittgenstein, L., Pears, D. F., McGuinness, B. y Russell, B. (2001). *Tractatus Logico Philosophicus* Routledge Classics.

Wittgenstein, L. (1980). *Piensiери Diversi*. (Nyman, H., Wright, G. H. y Ranchetti, M., trad). Adelphi Edizioni. (Obra Original publicada en 1970)

Zhang, R. (2022). Developing morphological knowledge with online corpora in an ESL vocabulary classroom. *Frontiers in Psychology*, 13, 4543. <https://doi.org/10.3389/fpsyg.2022.927636>

