



Fondecyt
Fondo Nacional de Desarrollo
Científico y Tecnológico

Universidad de Concepción
Dirección de Postgrado
Facultad de Humanidades y Arte
Programa de Doctorado en Lingüística

Análisis de colocaciones léxicas simples con verbos de apoyo y con verbos de significado pleno¹⁻²

TESIS PARA OPTAR AL GRADO DE
DOCTOR EN LINGÜÍSTICA



DIRECTORA DE LA TESIS: Dra. Anita Ferreira Cabrera

CANDIDATA: Lorena Blanco San Martín

CONCEPCIÓN, agosto de 2020.

¹ La realización de esta tesis doctoral fue posible gracias al financiamiento de la Comisión Nacional de Ciencia y Tecnología (CONICYT). CONICYT-PCHA/Doctorado Nacional Becas Chile/año 2015 – Folio 21151496.

² Esta tesis se enmarca en el proyecto FONDECYT N° 1180974 “Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua”.

RESUMEN

La producción de colocaciones léxicas simples verbo-nominales facilita la adquisición y el aprendizaje de una lengua extranjera. En este contexto, la presente tesis doctoral, investigación con un diseño no experimental de tipo descriptivo transversal que utiliza un enfoque de análisis de datos mixtos, tiene como objetivos generales: 1) Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE³ y 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1. Para ello, se han llevado a cabo dos estudios. En primer lugar, un estudio de frecuencia colocacional en un sub-corpus del corpus CAELE (Ferreira, 2014-2018) y, en segundo lugar, un estudio comparativo, que ha permitido establecer diferencias y similitudes entre ambos niveles de competencia del ELE.

En cuanto a la metodología, la investigación se ha sustentado en los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico. El procesamiento y análisis de datos se realizó en el *software UAM Corpus Tool* que permitió determinar las frecuencias de las colocaciones léxicas simples verbo-nominales. Asimismo, se utilizaron pruebas estadísticas para responder a los objetivos e hipótesis postuladas.

Los principales resultados indicaron que, por una parte, no hubo diferencia significativa en la cantidad de colocaciones producidas por los aprendientes de niveles de competencia del ELE A2 y B1. Por otra parte, sí se presentó cierta variación léxica en la selección de los tipos de bases y colocados.

Esta tesis aporta un listado de colocaciones léxicas simples verbo-nominales frecuentes en el ELE, las cuales mantienen estrecha relación con la lista de nociones específicas y nociones generales expuestas en el Plan Curricular del Instituto Cervantes (2006), entidad que se rige por los estándares del Marco Común Europeo de Referencia⁴ para las Lenguas (Consejo de Europa, 2001). Además, estas colocaciones tienen aplicabilidad dentro del campo de la

³ Corpus de Aprendientes de Español como Lengua Extranjera.

⁴ De aquí en adelante MCER.

enseñanza y el aprendizaje del ELE a través de distintas metodologías y tareas. También contribuyen a la descripción de la competencia comunicativa del ELE, puesto que han sido identificadas en un corpus de aprendientes electrónico que refleja la lengua en uso.

Palabras clave: Colocaciones léxicas simple verbo-nominales; colocaciones léxicas simples con verbos de apoyo; colocaciones léxicas simples con verbos de significado pleno; Corpus de Aprendientes de Lenguas en Formato Electrónico; CAELE y Español como Lengua Extranjera.



TABLA DE CONTENIDOS

Índice	Página
INTRODUCCIÓN	1
PLANTEAMIENTO DEL PROBLEMA	4
CAPÍTULO I. FUNDAMENTOS TEÓRICOS	9
1. Adquisición de una segunda lengua	9
1.1. Modelo cognitivo de ASL	9
1.2. Teoría sociocultural	10
1.3. Hipótesis de Krashen	10
2. Adquisición del léxico	11
3. La disciplina de la fraseología	13
3.1. Breve historia	13
3.2. Unidades fraseológicas y sus características	14
3.3. Enfoques fraseológicos	17
3.3.1. Enfoque combinatorio	17
3.3.2. Enfoque composicional	18
4. Las colocaciones	19
4.1. Historias de las colocaciones	19
4.2. Definiciones en torno al concepto de colocación	21
4.3. Características de las colocaciones	25
4.4. Niveles de las colocaciones	27
4.5. Colocaciones léxicas simples y complejas	25
4.6. Tipos de colocaciones verbo-nominales	29
5. Lingüística de Corpus	32
5.1. Corpus de aprendientes en formato electrónico	34
5.1.1. Corpus CAELE	37
5.1.2. Investigaciones a partir del Corpus CAELE	38
CAPÍTULO II. METODOLOGÍA DE LA INVESTIGACIÓN	41
6. Pregunta de investigación	41
6.1. Hipótesis	41
6.2. Objetivos de investigación	43
6.2.1. Objetivos generales	43
6.2.2. Objetivos específicos	43
6.3. Estudio 1: Frecuencia colocacional en el sub-corpus CAELE	44
6.3.1. Procedimientos metodológicos	45

6.3.2. Sub-corpus del corpus CAELE	46
6.3.3. Rasgos de los textos	47
6.3.4. Recolección de los datos	50
6.3.5. Descripción de los aprendices de ELE	50
6.4. Estudio 2: Comparación de la producción colocacional	51
6.4.1. Procedimientos metodológicos	53
6.4.2. Corpus de hablantes nativos de español	53
6.4.3. Sub-corpus del español	54
6.4.4 Rasgos de los textos	54
6.4.5. Recolección del sub-corpus	55
6.4.6. Descripción de los hablantes nativos de español	55
6.5. Procesamiento de los sub-corpus lingüísticos	56
6.5.1. Software UAM Corpus Tool 3.2.	56
6.6. Taxonomía para la anotación de colocaciones léxicas	57
6.7. Criterios para la identificación de colocaciones léxicas simples verbo-nominales	59
6.8. Revisión textual por parte de otros etiquetadores	61

CAPÍTULO III. RESULTADOS Y DISCUSIÓN **62**

7. Resultados Estudio 1: Frecuencia colocacional en el sub-corpus CAELE	63
7.1. Colocaciones léxicas verbo-nominales	64
7.2. Colocaciones léxicas más frecuentes con verbos de apoyo	64
7.3. Colocaciones léxicas más frecuentes con verbos plenos	65
7.3.1. Bases más frecuentes de las colocaciones léxicas verbo-nominales con verbos de significado pleno	66
7.3.2. Colocados más frecuentes de las colocaciones léxicas verbo-nominales con verbos de significado pleno	68
7.4. Prueba estadística <i>t-test</i>	69
7.5. Colocaciones léxicas por nivel de competencia comunicativa	70
7.5.1. Colocaciones léxicas verbo-nominales más frecuentes en el nivel de competencia A2	70
7.5.2. Colocaciones léxicas verbo-nominales más frecuentes en el nivel de competencia B1	70
7.6. Colocaciones léxicas verbo-nominales con verbo de apoyo más frecuentes por nivel de competencia comunicativa	71
7.6.1. Colocaciones léxicas verbo-nominales con verbos de apoyo más frecuentes en el nivel de competencia A2	71
7.6.2. Colocaciones léxicas verbo-nominales más frecuentes con verbos de apoyo en el nivel de competencia B1	73

7.7. Colocaciones léxicas con verbos de significado pleno por nivel de competencia comunicativa	74
7.7.1. Colocaciones léxicas verbo-nominales más frecuentes con verbos de significado pleno en el nivel de competencia A2	74
7.7.2. Colocaciones léxicas verbo-nominales más frecuentes con verbos de significado pleno en el nivel de competencia B1	75
7.8. Bases más frecuentes de colocaciones léxicas con verbos de apoyo por nivel de competencia comunicativa	77
7.8.1. Bases más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia A2	77
7.8.2. Bases más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia B1	78
7.9. Colocados más frecuentes de colocaciones léxicas con verbos de apoyo por nivel de competencia comunicativa	79
7.9.1. Colocados más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia A2	79
7.9.2. Colocados más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia B1	80
7.10. Bases más frecuentes de colocaciones léxicas con verbos de significado pleno por nivel de competencia comunicativa	81
7.10.1. Bases más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia A2	81
7.10.2. Bases más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia B1	82
7.11. Colocados más frecuentes de colocaciones léxicas con verbos de significado pleno por nivel de competencia comunicativa	83
7.11.1. Colocados más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia A2	83
7.11.2. Colocados más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia B1	84
8. Resultados Estudio 2: Comparación de la producción colocacional	85
8.1. Comparación de las colocaciones léxicas verbo-nominales en A2 y B1	85
8.2. Pruebas estadísticas <i>T-tests</i>	86
8.2.1. Colocaciones con verbos de apoyo en nivel A2 y B1	86
8.2.2. Colocaciones con verbos de significado pleno en nivel A2 y B1	87
8.3. Colocaciones en el sub-corpus de nativos de español	88

8.3.1. Colocaciones léxicas verbo-nominales	89
8.3.2. Colocaciones léxicas verbo-nominales con verbos de apoyo	89
8.3.3. Colocaciones léxicas con verbos de significado pleno	90
8.3.4. Bases de las colocaciones verbo-nominales con verbos de apoyo en el sub-corpus de nativos	92
8.3.5. Bases de las colocaciones verbo-nominales con verbos de significado pleno en el sub-corpus de nativos	93
8.3.6. Colocados de las colocaciones verbo-nominales con verbos de apoyo en el sub-corpus de nativos	94
8.3.7. Colocados de las colocaciones verbo-nominales con verbos de significado pleno en el sub-corpus de nativos	94
8.4. Colocaciones léxicas V+S en el sub-corpus CAELE y el corpus de nativos	95
8.5. Pruebas estadísticas ANOVA	96
8.5.1. Colocaciones léxicas simples verbo-nominales con verbos de apoyo	97
8.5.2. Colocaciones léxicas simples verbo-nominales con verbos de significado pleno	98
9. Discusión	100
9.1. Objetivo general 1) Determinar las colocaciones léxicas simples verbo-nominales más frecuentes de nivel A2 y B1 en el sub-corpus del corpus CAELE	100
9.2. Objetivo general 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acorde con el nivel A2 y B1	105
CAPÍTULO IV. CONCLUSIONES, LIMITACIONES Y PROYECCIONES DE ESTA INVESTIGACIÓN	112
10. Conclusiones	112
11. Limitaciones y proyecciones de la investigación	113
12. Referencias bibliográficas	115
13. Anexos	137
13.1. Etiquetado de colocaciones	137
13.2. Colocaciones léxicas verbo-nominales con verbos de significado pleno en el sub-corpus de ELE	139
13.3. Colocaciones léxicas con verbos de significado pleno de nivel A2	142
13.4. Colocaciones léxicas con verbos de significado pleno de nivel B1	145

Lista de tablas

Tabla 1. Clasificación de colocaciones léxicas simples (basada en Corpas, 1996 y Koike, 2001)	28
Tabla 2. Clasificación de colocaciones léxicas complejas (basada en Corpas, 1996 y Koike, 2001)	29
Tabla 3. Definición y operacionalización de las variables. Elaboración propia.	42
Tabla 4. Distribución del sub-corpus en los años de recopilación. Elaboración propia.	46
Tabla 5. Diseño y consideraciones para la construcción de un corpus de aprendientes (adaptada y traducida por Ferreira, Elejalde y Vine (2014), de Tono (2003)	47
Tabla 6. Distribución de tareas de escritura por tema textual. Elaboración propia.	48
Tabla 7. Número de palabras y <i>tokens</i> del sub-corpus de aprendientes. Elaboración propia.	49
Tabla 8. Distribución de los textos según lengua materna. Elaboración propia.	49
Tabla 9. Género de los aprendientes de ELE. Elaboración propia.	50
Tabla 10. Distribución de los aprendientes durante los años de recopilación del sub-corpus. Elaboración propia.	51
Tabla 11. Diseño y consideraciones para la construcción de un corpus de aprendientes (adaptada y traducida por Ferreira, Elejalde y Vine (2014), de Tono (2003).	54
Tabla 12. Número de palabras y <i>tokens</i> del corpus de nativos del español. Elaboración propia.	55
Tabla 13. Género de los hablantes nativos. Elaboración propia.	56
Tabla 14. Colocaciones léxicas V+S con verbos de apoyo en el sub-corpus CAELE. Elaboración propia.	65
Tabla 15. Colocaciones léxicas V+S con verbos de significado pleno en el sub-corpus CAELE. Elaboración propia.	66
Tabla 16. Bases en colocaciones léxicas simples V+S. Elaboración propia.	67

Tabla 17. T-test para grupos pareados en relación a tipos de colocaciones léxicas verbo-nominales. Elaboración propia.	69
Tabla 18. Colocaciones con verbos de apoyo en nivel A2. Elaboración propia.	72
Tabla 19. Colocaciones con verbos de apoyo en nivel B1. Elaboración propia.	73
Tabla 20. Colocaciones con verbos plenos en nivel A2. Elaboración propia.	74
Tabla 21. Colocaciones con verbos plenos en nivel B1. Elaboración propia.	76
Tabla 22. <i>T-test</i> para grupos diferenciados en relación a colocaciones con verbos de apoyo. Elaboración propia.	87
Tabla 23. <i>T-test</i> para grupos diferenciados en relación a colocaciones con verbos plenos.	87
Tabla 24. Colocaciones léxicas simples con verbos de apoyo de hablantes nativos de español. Elaboración propia.	90
Tabla 25. Colocaciones léxicas simples con verbos de significado pleno de hablantes nativos de español. Elaboración propia.	91
Tabla 26. Resumen de los datos de colocaciones léxicas-simples con verbos de apoyo. Elaboración propia.	97
Tabla 27. Resumen de ANOVA de colocaciones léxicas-simples con verbos de apoyo. Elaboración propia.	97
Tabla 28. Resumen de los datos de colocaciones léxicas-simples con verbos de significado pleno Elaboración propia.	98
Tabla 29. Resumen de ANOVA de colocaciones léxicas-simples con verbos de apoyo Elaboración propia.	99

Lista de figuras

Figura 1. Taxonomía estudio de frecuencia colocacional	58
Figura 2. Aplicación de taxonomía	59

Lista de gráficos

Gráfico 1. Colocaciones léxicas V+S en el sub-corpus CAELE. Elaboración propia.	64
Gráfico 2. Colocados en colocaciones léxicas simples V+S. Elaboración propia.	68
Gráfico 3. Colocaciones léxicas verbo-nominales en nivel A2. Elaboración propia.	70
Gráfico 4. Colocaciones léxicas verbo-nominales en nivel B1. Elaboración propia.	71
Gráfico 5. Bases con verbos de apoyo en nivel A2. Elaboración propia.	77
Gráfico 6. Bases con verbos de apoyo en nivel B1. Elaboración propia.	78
Gráfico 7. Colocados con verbos de apoyo en nivel A2. Elaboración propia.	79
Gráfico 8. Colocados con verbos de apoyo en nivel B1. Elaboración propia.	80
Gráfico 9. Bases con verbos plenos en nivel A2. Elaboración propia.	81
Gráfico 10. Bases con verbos plenos en nivel B1. Elaboración propia.	82
Gráfico 11. Colocados con verbos plenos en nivel A2. Elaboración propia.	83
Gráfico 12. Colocados con verbos plenos en nivel B1.	84
Gráfico 13. Colocaciones léxicas verbo-nominales en niveles de competencia lingüística. Elaboración propia.	86
Gráfico 14. Colocaciones léxicas V+S en el sub-corpus de nativos. Elaboración propia.	89
Gráfico 15. Bases con verbos de apoyo en el sub-corpus de nativos. Elaboración propia.	92
Gráfico 16. Bases con verbos plenos en el sub-corpus de nativos. Elaboración propia.	93
Gráfico 17. Colocados con verbos de apoyo en el sub-corpus de nativos. Elaboración propia.	94
Gráfico 18. Colocados con verbos plenos en el sub-corpus de nativos. Elaboración propia.	95
Gráfico 19. Comparación de colocaciones léxicas entre el sub-corpus CAELE y el sub-corpus de nativos. Elaboración propia.	96

DEDICATORIA

A Alonsito.



AGRADECIMIENTOS

En primer lugar, agradezco a Dios por el don de la vida, la oportunidad de estudiar y por darme alas para volar, las cuales me han permitido conquistar anhelos y cumplir sueños.

En segundo lugar, agradezco a mi familia, especialmente a María Elena, mi madre, y a Alonso, mi hijo. Gracias por su amor durante estos años de estudio.

En tercer lugar, doy las gracias a la Dra. Anita Ferreira por su infinita paciencia, excelente disposición y apoyo durante esta etapa de formación. Agradezco su rigurosa guía metodológica sin la cual el desarrollo de esta tesis no hubiese sido posible.

En cuarto lugar, agradezco al Programa ELE-UDEC, fundado por la Dra. Ferreira. Allí impartí docencia y logré observar distintas problemáticas de la lengua que me llevaron a definir los objetivos de investigación de esta tesis. Asimismo, agradezco al Proyecto FONDECYT N° 1180974 “Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua” por proveerme distintas herramientas metodológicas.

En quinto lugar, agradezco a mis amigos de la vida por su apoyo incondicional durante este proceso. Gracias también a mis colegas del Programa de Doctorado por el compañerismo, la discusión académica y la amistad.

Finalmente, agradezco a la Comisión Nacional de Ciencia y Tecnología, por financiar estos 5 años de estudio.

INTRODUCCIÓN

Una unidad léxica puede ser definida como una unidad de la lengua con valor referencial que forma parte del léxico de los hablantes, quienes pertenecen a una comunidad lingüística particular (Cid, 2012). La literatura científica reconoce que “los hablantes adquieren de forma natural e inconsciente las unidades léxicas complejas de su idioma; pero, para quienes aprenden una segunda lengua, estas unidades dificultan frecuentemente el desarrollo amplio de su competencia lingüística” (Soto, 2012, p. 3). A partir de esta dificultad, el estudio de una segunda lengua genera complejidad y suscita el interés por parte de distintas disciplinas.

Los estudios colocacionales en el ámbito de ELE datan desde la década del 80 aproximadamente. El primer investigador que introdujo el término colocación fue Seco (1978). Este fenómeno de estudio es parte esencial en la adquisición de L2. Con su exploración, identificación y posterior análisis se logra caracterizar la competencia colocacional de hablantes extranjeros respecto de una segunda lengua. Asimismo, constituye un principio esencial para la adquisición del léxico, el cual es almacenado en el lexicón mental de los aprendientes (Higueras, 2004). Desde esta perspectiva, el estudio de las colocaciones en la adquisición de una Lengua Extranjera (LE) o segunda lengua (L2)¹ se encuentra cada vez más vigente en las comunidades de investigación. Este fenómeno aparece tanto en el lenguaje cotidiano como en contextos más específicos. Para algunos investigadores, es requisito esencial para el proceso de enseñanza/aprendizaje de una LE que el aprendiz domine las colocaciones (Lewis, 2000a) para desarrollar su competencia comunicativa a través de la recuperación y aprendizaje de los bloques léxicos que conforman estas estructuras.

Las colocaciones léxicas simples con verbos de apoyo y verbos de significado pleno constituyen un tipo de colocación elemental en la producción oral y escrita por parte de los aprendientes de ELE, puesto que estas estructuras están

¹ A lo largo de la tesis, se usan indistintamente los términos segunda lengua (L2), lengua extranjera (LE), lengua meta y lengua objeto para referirnos a una lengua que se adquiere con posterioridad a la L1 o lengua materna (LM) (Pastor, 2004).

conformadas por dos clases léxicas variables y altamente productivas en el español, verbo y sustantivo. En este sentido, actualmente, en el ámbito de ELE, no hay estudios concluyentes a) que problematicen tanto las descripciones como las frecuencias de colocaciones léxicas simples (Wen, 2019), específicamente, en las combinaciones verbo-nominales conformadas por verbos de apoyo y por verbos de significado pleno en los niveles A2 y B1 del ELE (Pérez, 2014; De Alba Quiñones, 2014; Uriel, 2014; Lozano y Rodríguez, 2014, De Miguel, 2008; Sidoti, 2014 y Molero y Salazar, 2013) y b) que estén basadas en corpus especializados, exactamente, en corpus de aprendientes en formato digital. En consecuencia, surge la necesidad de abordar la naturaleza de este tipo de colocaciones en tareas de escritura, producidas en un contexto auténtico de aprendizaje.

La presente investigación se enmarca en la interdisciplina de la Lingüística Aplicada (LA), concretamente, en a) la adquisición y aprendizaje en el ámbito de Español como Lengua Extranjera (ELE), b) la enseñanza de segundas lenguas, centrada en la competencia comunicativa del aprendiente, y c) la metodología de la Lingüística de Corpus (LC). Esta metodología “se focaliza en el estudio de la lengua en contextos de uso real para observar los fenómenos ocurrentes” (Ferreira, Elejalde y Vine, 2014, p.387) utilizando principalmente los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico y el *software UAM Corpus Tool*. Los objetivos generales son 1) Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE² y 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1.

En cuanto a su metodología, la investigación utilizó un diseño no experimental de tipo descriptivo transversal para identificar y analizar los usos presentes en las colocaciones léxicas simples verbo-nominales de los aprendientes de ELE. Como parte de los procedimientos, los resultados fueron comparados entre los niveles

² Corpus de Aprendientes de Español como Lengua Extranjera.

objeto de estudio y luego con un sub-corpus de hablantes nativos de español como lengua materna (LM)³, a través de la aplicación de pruebas estadísticas.

Respecto a la organización, esta tesis considera las siguientes partes. En primer lugar, se presenta la formulación del problema. En segundo lugar, se presenta los principales planteamientos y sustentos teóricos. En tercer lugar, se expone la metodología de investigación. En cuarto lugar, se presenta los resultados y la discusión. En quinto y último lugar se exponen las conclusiones, limitaciones y proyecciones de esta investigación.

Finalmente, esta tesis doctoral, incluida la base metodológica y el corpus, se desarrolló en el ámbito de la investigación del proyecto FONDECYT N° 1180974 (Ferreira, 2018-2021) “Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua”.



³ El sub-corpus de español pertenece al corpus de español recolectado por Kloss (2018) en el contexto de su tesis doctoral.

PLANTEAMIENTO DEL PROBLEMA

El léxico constituye un aspecto fundamental dentro del campo de la Lingüística Aplicada. Su adquisición permite que los aprendientes conozcan, almacenen, dominen y produzcan nuevas palabras y estructuras de la lengua meta desarrollando no solo la competencia léxica, sino también la competencia comunicativa e intercultural mediante la práctica de sus habilidades lingüísticas.

Desde la perspectiva de la Lingüística de Corpus, los corpus especializados, específicamente, los corpus de aprendientes en formato digital pueden ser definidos como la colección de datos lingüísticos auténticos (textos orales o escritos) en que se constata el uso de la lengua objeto de estudio (segunda lengua, L2, o lengua extranjera, LE) (Granger, 2003; 2004). Constituyen una valiosa metodología que facilita, sin duda, el procesamiento de un número no menor de textos, los cuales, si fueran procesados manualmente, implicaría una inversión mayor de tiempo para el investigador. A través de las investigaciones, se puede conocer qué aspectos lingüísticos se aprenden antes que otros, cuáles son más difíciles de aprender y las diferencias con los hablantes nativos (HN), por medios de análisis contrastivos (Granger, 2015; Lozano, 2015). Por lo tanto, este tipo de corpus es concebido como un valioso recurso para la investigación en el ámbito de la Adquisición de segundas lenguas (ASL) y en la enseñanza de L2

Trabajar con corpus lingüísticos permite mostrar la realidad lingüística de una comunidad determinada, corroborar o rechazar hipótesis, teorías e ideas y si la muestra es representativa, y establecer bajo parámetros objetivos reglas que permitan explicar el fenómeno que se estudia (Hunston, 2006; Pérez-Ávila 2006; Cruz Piñol, 2012); sin embargo, en el caso de los corpus lingüísticos de lenguas extranjeras, al ser muestras auténticas generadas en diferentes ambientes de aprendizaje se observa el estado actual de la interlengua y se comprueba el empleo de una lengua determinada para conocer los errores o aciertos sobre alguna problemática lingüística (Nesselhauf, 2004). Desde esta óptica, la influencia del corpus lingüístico es fundamental cuando se estudia el léxico y sus

usos, especialmente, porque la base de este tipo de estudios son las colocaciones y los grupos léxicos (Nattinger y De Carrico, 1992; Schmitt, 2004).

El estudio de los corpus de aprendientes en formato digital permite acceder al léxico y contribuir con evidencia empírica en la descripción de los procesos de adquisición de la competencia lingüística ya sea en una LM o en una LE. En este sentido, la competencia léxica es altamente relevante en el proceso de adquisición, dado que las unidades léxicas tanto simples como complejas portan el significado y permiten la comunicación, la interacción y la negociación entre hablantes de una lengua extranjera y hablantes nativos (Long, 1980). Si bien esta competencia puede ser estudiada a través de unidades de vocabulario individualizadas, también, a través de las colocaciones, estructuras frecuentes fijadas por norma y presentes en una determinada lengua.

Para Corpas (1996), las colocaciones son sintagmas libres, generados a partir de reglas que presentan cierto grado de restricción combinatoria, determinadas por el uso. Su naturaleza puede ser tanto gramatical como léxica y su producción varía de simple a compleja, por lo que aprender una palabra significa también aprender con qué otras concurre (Nation, 2001). Las colocaciones también son definidas como elementos indispensables del discurso natural, y su aprendizaje es necesario para poder comunicar con fluidez, corrección y precisión. Como señalan Alonso y Kalan (2004):

La selección apropiada de colocaciones es un “visado lingüístico” que permite a cualquier estudiante de L2 acercarse a la competencia lingüística de un nativo. Le posibilita el que su producción sea más fluida, correcta, precisa y compleja. Para el estudiante de ELE, elegir la colocación adecuada supone producir un texto mucho más natural, mucho más fluido, mucho más cercano al que un nativo produciría (p. 821).

Se ha seleccionado como objeto de estudio las colocaciones léxicas simples con verbos de significado pleno y verbos de apoyo (colocaciones verbo-nominales), dado que existe poca evidencia que describa y compare por niveles de

competencia del ELE la producción de estas combinaciones léxicas en corpus de aprendientes en formato electrónico. Actualmente, el Plan curricular del Instituto Cervantes (2006) no incluye explícitamente este tipo de colocaciones en su lista de nociones generales ni tampoco en su lista de nociones específicas. Además, en palabras de Koike (2005) “en las colocaciones simples, las combinaciones sustantivo-verbo son las más representativas en su número” (pp.183). En virtud de estos argumentos, es necesario la realización de esta investigación, cuyos resultados contribuirán a la descripción de la competencia colocacional de los aprendientes de niveles iniciales e intermedios del ELE.

Según su naturaleza, este tipo de combinaciones está conformado por a) colocaciones léxicas simples con verbos de apoyo o construcciones con verbo soporte (Alonso, 1997), aquellas en que los verbos tienen una naturaleza deslexicalizada y carecen de autonomía o valor semántico cuando son combinadas con nombres que denotan eventos y b) colocaciones léxicas simples con verbos de significado pleno. En ambas combinaciones el verbo actúa como colocado y el sustantivo como base, por lo que este último se encarga de realizar la elección léxico-semántica.

Higueras (2004) enfatiza que el aspecto colocacional cobra especial importancia en la adquisición/aprendizaje de una lengua segunda, puesto que las restricciones de combinatoria que presentan numerosas palabras suelen pasar desapercibidas entre la gran mayoría de estudiantes de una lengua extranjera. Entonces, aprender colocaciones no solo facilita la adquisición de nuevo vocabulario, es también una manera de enriquecer el vocabulario ya adquirido (Higueras, 2006) y presente en el lexicón mental.

El problema de las colocaciones léxicas verbo-nominales que se está estudiando apunta a la producción de estas combinaciones en el nivel inicial y en el nivel intermedio de competencia lingüística del español, puesto que los aprendientes⁴ de nivel B1 deberían producir mayor cantidad de colocaciones que los

⁴ De aquí en adelante, se utilizará el término aprendiente(s) en vez de aprendiz o aprendices. Este último solo aparecerá cuando sea extraído textualmente de alguna investigación.

aprendientes de nivel intermedio. Actualmente, este tipo de colocaciones léxicas en estos niveles ha sido poco estudiado. Si bien existen valiosos diccionarios de colocaciones, existe la necesidad de describir el comportamiento de estas unidades léxicas en corpus lingüísticos producidos por aprendices de ELE. Al respecto, González (2006) plantea la necesidad de investigar el tema, porque, por ejemplo, en el ámbito de los negocios contribuye a la enseñanza y aprendizaje de ELE. Por su parte, Fernández (2005) propone profundizar en la investigación de las colocaciones, expresiones muy abandonadas hasta el momento. La falta de investigaciones que aborden esta problemática lingüística tiene como consecuencia que las colocaciones sean poco trabajadas en el aula de ELE. Por lo tanto, se reclama obras específicas y nuevas metodologías para la enseñanza de unidades fraseológicas (Higuera, 1997; Prieto, Mosqueira y Vásquez, 2009; Ruiz, 2000; Torjada, 2000), en las cuales se incluyan las combinaciones objeto de estudio.

No existen estudios concluyentes sobre frecuencias ni descripciones de colocaciones léxicas verbo-nominales por niveles de competencia comunicativa en ELE. Así lo demuestran los estudios realizados por Pérez (2014), De Alba Quiñones (2014), Uriel (2014), Lozano y Rodríguez (2014), De Miguel (2008), Sidoti (2014) y Molero y Salazar (2013). Dichas investigaciones consideran la competencia léxica como esencial en el aprendizaje de una lengua extranjera. En el ámbito de la didáctica, Liu (2010) señala que en los materiales centrados en la enseñanza de las colocaciones existe una falta de análisis crítico sobre la definición y naturaleza de las colocaciones así como de la metodología para la enseñanza de las mismas. Desde esta perspectiva, Zarco (1998) enfatiza la importancia de este tipo de construcciones léxicas en traducción y en la didáctica de las segundas lenguas, ya que su naturaleza semilexicalizada o perifrástica, locucional o colocacional es una indudable fuente de problemas lingüísticos.

En un contexto metodológico, la exploración de un corpus permite la identificación y posteriormente el análisis de estructuras colocacionales utilizando *softwares* computacionales para la obtención de resultados de forma semiautomática y

automática. En consecuencia con lo anterior, los corpus lingüísticos de aprendientes constituyen datos auténticos que permiten evidenciar la competencia lingüística de los aprendientes de segundas lenguas. En esta competencia se incluye la competencia colocacional, una sub-competencia de la competencia léxica, que debe ser descrita parcial o totalmente en los distintos niveles de competencia comunicativa de los aprendientes de ELE.

A la luz de los planteamientos expuestos, esta tesis doctoral se ha propuesto responder la siguiente pregunta de investigación: ¿Varían las colocaciones léxicas simples con verbos de apoyo y con verbos de significado pleno (colocaciones verbo-nominales) producidas por los aprendientes de niveles del ELE A2 y B1? Para ello, se han planteado dos objetivos generales: 1) Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE y 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1.



CAPÍTULO I. FUNDAMENTOS TEÓRICOS

A continuación, se presentan los principales fundamentos teóricos en los que se circunscribe la problemática de estudio.

1. Adquisición de Segundas Lenguas

En las últimas décadas, se ha investigado la adquisición y aprendizaje de segundas lenguas con la finalidad de descubrir cómo el ser humano llega a aprender otra lengua. Es así como los estudios sobre la ASL se enmarcan en el ámbito interdisciplinar de la LA intentan dar explicación a este fenómeno complejo. En el proceso de ASL, surgen dos paradigmas que contribuyen con sus teorías a este campo de investigación. Por una parte, el paradigma basado en las teorías cognitivas provenientes de la psicología y la lingüística. Por otra parte, el paradigma basado en la teoría sociocultural.



1.1. Modelo cognitivo de ASL

Este modelo se centra en los procesos cognitivos que involucran el aprendizaje y uso de una segunda lengua. Recibe el aporte de los modelos computacionales que facilitan el aprendizaje de una lengua meta. Lamy y Hampel (2007) señalan que la adquisición es el producto del procesamiento del *input*⁵ y del *ouput*⁶. El *input* corresponde a las exposiciones del aprendiente a la lengua meta, mientras que el *ouput*, a la producción realizada por cada aprendiente. *Input* y *ouput* están mediados por la interacción comunicativa. En este contexto, Lamy y Hampel (2007) plantean la importancia de la interacción entre los aprendientes y otros hablantes para el desarrollo del lenguaje.

⁵ Muestras de lengua meta, orales o escritas, que el aprendiente encuentra durante su proceso de aprendizaje y a partir de las cuales puede realizar ese proceso. (https://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/indice.htm#i).

⁶ Producción oral o escrita por parte de un aprendiente de LE.

1.2. Teoría sociocultural

El aporte de la teoría sociocultural desarrollada por el psicólogo ruso Lev Vygotsky (1896-1934) ha sido de gran relevancia en el contexto educativo por el impacto que ha generado en los diseños curriculares y didácticos de la enseñanza. Esta teoría considera el concepto de mediación. Warschauer (2005) señala que: “the notion that all human activity is mediated by tools or signs” (pp. 41)⁷, lo anterior implica que todos los recursos o herramientas tales como la tecnología o el mismo lenguaje actúan como mediadores en el sentido que facilitan el proceso de adquisición y los procesos de conducta o de comportamiento del ser humano.

El valor de la teoría sociocultural radica en el contexto social de los aprendientes, quienes a través de la interacción desarrollan sus procesos cognitivos en post del desarrollo de la competencia comunicativa.

Existen otros fundamentos teóricos que describen el proceso de adquisición de una segunda lengua. El Modelo del Monitor postulado por Krashen (1982) se encuentra basado en cinco hipótesis.



1.3. Hipótesis de Krashen

1. La hipótesis de la adquisición y el aprendizaje: adquisición y aprendizaje se configuran como dos procesos diferentes. El aprendizaje es consciente y la adquisición, inconsciente.
2. La hipótesis del orden natural: la enseñanza puede facilitar el aprendizaje de los aspectos lingüísticos siempre y cuando estos sean en un orden natural.
3. La hipótesis del monitor: el aprendiz es consciente cuando usa su lengua. Constantemente edita el *output* para corregir la comunicación y lograr que esta sea efectiva.

⁷ “La noción de que toda la actividad humana está mediada por herramientas o signos”.

4. La hipótesis del filtro afectivo: las motivaciones y la actitud del aprendiente influyen positiva o negativamente en el proceso de adquisición o aprendizaje de una segunda lengua.
5. La hipótesis del *input*: el aprendiz aprende la lengua cuando comprende el mensaje del *input*, es decir, cuando hay un *input* comprensible.

Las hipótesis anteriores fundamentan el proceso de ASL. Asimismo, abren paso al desarrollo del léxico en los aprendientes de una LE.

2. Adquisición del léxico

En primer lugar, el proceso de adquisición del léxico en una segunda lengua ha sido ampliamente estudiado por su importancia al momento de establecer comunicación. Se puede constatar algunas diferencias básicas en comparación con la adquisición de la primera lengua (Neuner-Anfindsen, 2005):

- El aprendiz de una segunda lengua se encuentra en una fase de evolución fisiológica y cognitiva más avanzada.
- El *input* que recibe un aprendiz de una segunda lengua es muy distinto al que recibe un niño que aprende su lengua materna.
- El aprendiz de una segunda lengua ya ha pasado por el aprendizaje de una lengua, por lo que ya ha superado las primeras fases de la evolución léxica.

Siguiendo los postulados de Neuner-Anfindsen (2005), el léxico constituye un proceso distinto cuando el aprendiz se expone a *input* de una segunda lengua, dado que se halla en una fase cognitiva más avanzada al haber sido ya expuesto a los *inputs* de su lengua materna.

En segundo lugar, el aprendizaje de cada ítem léxico es un aprendizaje continuo

formado por diferentes procesos graduales, en los cuales las capacidades cognitivas y las habilidades lingüísticas se desarrollan gradualmente. Desde esta perspectiva, Izquierdo (2004:46-47) considera que el aprendizaje de una unidad léxica engloba una serie de sub-aprendizajes:

- a) Reconocer la unidad léxica cuando se oye y saber pronunciarla.
- b) Reconocer la forma escrita de la unidad léxica y ser capaz de escribirla.
- c) Reconocer la morfología de la unidad léxica, es decir, los morfemas que la forman, relacionar dichas partes con su significado, así como ser capaz de formar la unidad léxica utilizando los morfemas correctos.
- d) Reconocer las diferentes acepciones o significados y ser capaz de producir la unidad léxica para expresar su significado según el contexto.
- e) Reconocer su categoría gramatical.
- f) Conocer las estructuras sintácticas en las que puede aparecer y sus restricciones.

Los sub-aprendizajes señalados por Izquierdo (2004) se complementan con la idea de que una sola unidad va transformándose en la medida que los procesos cognitivos del aprendiz de una LE aumentan y se tornan cada vez más complejos. En la medida que el aprendiente se vea enfrentado a tareas posibilitadoras, significativas e interactivas, su aprendizaje no solo será explícito; sino también, implícito, esto es, un conocimiento de carácter procedimental e inconsciente que afectará positivamente tanto la comprensión como la producción de distintas

formas lingüísticas, por ejemplo, frases, oraciones, locuciones, colocaciones, entre otras.

3. La disciplina de la fraseología

3.1. Breve historia

Es indudable que existe una amplia variedad conceptual referida a la unidad fraseológica. Algunos términos encontrados en la literatura son: “expresión pluriverbal”, “unidad pluriverbal lexicalizada y habitualizada”, “unidad léxica pluriverbal”, “expresión fija” o “fraseologismo.” Esto ha generado un problema terminológico, caracterizado por la inestabilidad.

Formalmente la fraseología se constituyó como disciplina por los lingüistas soviéticos durante el siglo XX (1940). Vinogradov (1947) estudió el funcionamiento de estas unidades en el idioma ruso y con ello estableció las bases teóricas. Clasificó estas combinaciones de palabras en: 1) adherencias fraseológicas (estructuras con un alto grado de cohesión semántica), 2) unidades fraseológicas (unidades con significado unitario, pero identificable por alguno de los elementos que la conforman), 3) combinaciones fraseológicas (unidades que adquieren un determinado valor en relación al contexto en el cual se usen o aparezcan) y 4) expresiones fraseológicas (estructuras fijas, tales como refranes, frases clichés, modismos y proverbios).

Estos postulados tuvieron alcance en Europa. Por una parte, en Alemania destacan investigadores como Hausmann (1979), Wotjak (1992). Dobrovolskij (1992), Burger (1973), Burger, Buhofer y Sialm (1986) y Fleischer (1982). Este último logró consolidar estos fundamentos en su obra “Manual sobre la fraseología del alemán contemporáneo”, considerado como un clásico en esta disciplina. Por otra parte, Europa occidental también desarrolló estudios en torno a esta teoría fraseológica.

Julio Casares (1950) fue el propulsor en el ámbito de la lengua española, ya que propuso una taxonomía en torno a este fenómeno lingüístico, la cual tuvo un gran impacto para los investigadores venideros, entre los que destacan Coseriu (1977), quien estableció diferencias entre discurso repetido y técnica libre del discurso, Zuluaga (1980), Tristán (1988), Ettinger (1992), Corpas (1996), Ruiz (1997), entre otros. Corpas (1996) fue quien se encargó de replantear el estado del arte de las unidades fraseológicas afirmando que estas unidades se dividen en: colocaciones, locuciones y enunciados fraseológicos, y que se corresponden con elementos de la lengua. Es así como las colocaciones se encuentran fijadas por la norma, las locuciones, por el sistema, mientras que los enunciados tienen cabida en el habla. En este sentido, destaca la precisión de que los dos primeros elementos se restringen al plano de las frases o los sintagmas, a diferencia de los enunciados que traspasan ese umbral y se posicionan dentro del discurso y la interacción que mantienen los hablantes. Actualmente, en LM o LE distintas disciplinas, tales como la Lingüística Cognitiva y el Análisis del Discurso ofrecen más posibilidades para un tratamiento teórico adecuado de este fenómeno lingüístico.

3.2. Unidades fraseológicas y sus características

Las características principales que definen una unidad fraseológica son la fijación y la idiomatidad. No obstante, también se presentan otras condiciones tales como la lexicalización, la frecuencia de coaparición y la pluriverbalidad.

1) La fijación puede ser definida como “la propiedad que tiene ciertas expresiones de ser reproducidas en el hablar como combinaciones previamente hechas” (Zuluaga, 1990, p. 230). En este escenario, la fijación tiene mayor protagonismo en las habilidades comunicativas productivas. Thun (1978) señala que esta característica se presenta a través de la “fijación interna” y de la “fijación externa”. La primera engloba la realidad material, es decir, la realización fonética fija e inalterable de sus componentes (“a ciencia cierta”/* “a cierta ciencia”; “de armas tomar”/* “de tomar armas”). Esto hace que la elección de las categorías léxicas no

sea libre, sino restrictiva (Zuluaga, 1990). También este tipo de fijación permite expresar la invariabilidad de rasgos gramaticales, por ejemplo: género, número, tiempo, persona (quien mucho abarca poco aprieta”/* “quien mucho abarca poco apretó”, “tirar la toalla”/* “tirar las toallas”). La segunda abarca: la situación social, lo que quiere decir que las unidades fraseológicas se enmarcan en un contexto específico (fijación situacional); el uso de determinadas unidades lingüísticas frente a otras igualmente posibles teóricamente (fijación analítica); el empleo de unidades lingüísticas según el papel del hablante en el acto comunicativo (fijación pasemática); la preferencia de uso de ciertas unidades fraseológicas en la producción de textos (fijación posicional).

2) La idiomaticidad puede entenderse de diferentes maneras. Etimológicamente, se entiende como lo que es propio y particular de un determinado sistema lingüístico. Desde otro punto de vista, es el rasgo semántico característico de algunas combinaciones fijas, cuyo sentido no puede deducirse de la suma de los significados de los elementos constituyentes (Zuluaga, 1980, p. 121). A su vez, puede ser definida como una propiedad semántica que presenta ciertas unidades fraseológicas, por la cual el significado global de dicha unidad no es deducible del significado aislado de cada uno de sus elementos constitutivos (Fernando y Flavell, 1981). Para Zuluaga (1980):

la expresión idiomática es un signo complejo, pero no simultáneamente y desde el punto de vista funcional, un complejo de signos. Los componentes de esta no se comportan en ella como signos lingüísticos propiamente, sino más bien como componentes formales de un signo (p.123-124).

El significado de sus componentes no se entiende individualmente, sino en bloque. En términos de funcionalidad, se puede establecer que es un indicador potencial y no determinante, “se trata de una característica potencial y no esencial, de este tipo de unidades” (Corpas, 1996, p. 27), dado que no todas las unidades fraseológicas pueden ser consideradas idiomáticas. Asimismo, una palabra también puede calificar como tal.

3) Por lexicalización (especialización semántica) se ha de entender la variabilidad semántica que puede experimentar una unidad fraseológica (UF) una vez establecida en el uso, es decir, “una combinación de palabras se fija primero y después es cuando se convierte en una unidad potencial susceptible de modificar su significado” (Corpas, 1996, p. 26). Por ejemplo, en “tirar o arrojar la toalla” se evidencia la construcción de un significado unitario. Este significado varía. Puede referirse a una situación estrictamente deportiva. Por ejemplo. “El campeón tiró la toalla en el tercer asalto” o extrapolarse a una situación más amplia, por ejemplo. “Estaba estudiando, pero tiré la toalla”. En ambos ejemplos el valor de la expresión implica abandonar o renunciar a alguna actividad.

Conviene señalar que algunos elementos que componen la unidad fraseológica pueden sustituirse por otros. Esta característica parece oponerse a la de variación léxica, no obstante, como puntualiza Carneado (1985): “en ningún caso la modificación de un giro fraseológico lo altera si se realiza en aquellos puntos de la estructura que no cumplen una función diferenciadora” (p. 270-271). Además, una UF puede variar, pero esa variación ya está fijada. Ej.: “quedar todo en casa/ en familia/* en hogar”. Si se altera alguno de los componentes fijados en la combinación entraríamos ante el fenómeno de la desautomatización (Zuluaga, 1980) generalmente para producir efectos expresivos o estilísticos.

4) En lo que respecta a la coaparición, esta se expresa en frecuencia, por lo tanto, la aparición de los elementos en una muestra representativa oral o escrita constituye un indicador para determinar una unidad fraseológica. Las unidades prefabricadas se repiten en una comunidad de hablantes, en contextos o situaciones comunicativas. Esta coaparición sienta las bases para la institucionalización o convencionalización de las unidades léxicas.

5) No todas las unidades compuestas o estructuras combinadas califican como unidades fraseológicas. Algunas unidades se caracterizan por presentar una sintaxis libre.

3.3. Enfoques fraseológicos

El aporte de la teoría rusa ha sido fundamental en el estudio de las unidades fraseológicas al igual que los aportes de la teoría fraseológica basada en el análisis de corpus y los aportes de la tradición antropológica. Estos fundamentos configuran la base del enfoque combinatorio. Los estudios desarrollados por los lingüistas occidentales constituyen el núcleo teórico-práctico del enfoque composicional, que surge del enfoque Neo-Firthian, 'co-occurrence' tradition of lexicology (Hanks, 2000, p. 213)⁸, hacia 1970.

3.3.1. Enfoque combinatorio

Tal como su nombre lo indica, este enfoque combina los niveles de la lengua, específicamente, el léxico y el gramatical. Su finalidad apunta a identificar interna y externamente las UF. Para ello las distingue entre combinaciones léxicas y combinaciones sintácticas. Las tendencias que fundamentan el enfoque son:

- a) La teoría rusa clásica, base de la teoría fraseológica, es incomparable en su aplicación para el diseño de diccionarios. Uno de los autores que pertenece a esta corriente es Mel'čuk (1998, 1992, 1996).
- b) Siguiendo el modelo de la teoría anterior, la vertiente antropológica se centra en la dimensión cultural de la producción de las UF.
- c) La teoría fraseológica basada en el almacenamiento y procesamiento electrónico de corporas lingüísticas del inglés hablado y escrito. Su contribución radica en el proceso de enseñanza-aprendizaje, dado que se centra en la creación de diccionarios de UF a través de las herramientas informáticas.

⁸ Tradición de la lexicología de la co-ocurrencia.

3.3.2. Enfoque composicional

Este enfoque considera las construcciones fraseológicas como un todo, un conjunto indivisible al momento de generar significado. Sinclair (2002) asegura en relación al *idiom principle*⁹, que “los términos de una lengua tienden a unirse a otros para dar un significado nuevo” (p. 213). Por lo tanto, la característica fundamental de este enfoque es el análisis de frecuencias de coaparición de las unidades basándose en el análisis de corpus lingüísticos. Al respecto, Pawley (2001) describe los objetivos de este enfoque:

- a) Identificar en los córpora cualquier tipo de colocación, ya sean combinaciones libres o restringidas, y delimitar dónde terminan unas y comienzan las otras.
- b) Diseñar una tipología de combinaciones restringidas teniendo en cuenta factores estructurales y funcionales.
- c) Explicar la prevalencia y frecuencia de estas expresiones en cualquier tipo de discurso. (p. 125)

En resumen, este enfoque centra su estudio en las unidades fraseológicas en términos de colocabilidad, es decir, en qué expresiones aparecen con mayor y menor fuerza en el discurso analizado y en qué grado (Butler, 1998, p.16). Este último indicador se refiere a la atracción que se genera entre los elementos que componen la colocación.

4. Las colocaciones

4.1. Historia de las colocaciones

Su origen se remonta a las pesquisas realizadas a principio del siglo XX por el estructuralista Ferdinand de Saussure (Corpas, 1996). El francés Bally (1951) estudió algunas propiedades de las colocaciones, tal como lo es la coocurrencia

⁹ Principio idiomático.

léxica. Décadas después, Porzig (1950) comenzó estudios sobre algunas propiedades de estas combinaciones, pero fue Firth (1957), quien acuñó por primera vez el término en su artículo «Modes of Meaning»¹⁰, cuya finalidad era determinar el estilo literario de un autor. Postuló no solo la distinción entre dos tipos de colocaciones: usuales y restringidas, sino también que la coaparición de una unidad léxica en relación a otras palabras es la que define su significado. Sus hallazgos lo llevaron a introducir formalmente este objeto de investigación en la lingüística.

Halliday (1966) introduce el concepto de “lexicalidad” y enfatiza la diferencia que existe en el estudio colocacional cuando se estudia desde un enfoque gramatical o léxico. Por su parte, considera el análisis de las colocaciones desde un nivel léxico. En consonancia con esto, Sinclair (1966) insiste en el comportamiento diferente de una colocación a nivel léxico y gramatical apostando un paso más allá, por lo que descarta meramente la perspectiva gramatical. Además, introduce los términos *node*¹¹ y *collocates*¹², elementos que forman una colocación. Mientras la base es semánticamente independiente y determina una UL, el colocativo expresa un significado concreto. Ambos autores postulan que es el enfoque estadístico el que permite el análisis colocacional, gracias a la coocurrencia de los elementos léxicos (Halliday 1966; Sinclair, 1966). No obstante, para algunos investigadores, la coocurrencia frecuente no sería un indicador suficiente para determinar una colocación (Alonso, 1994; Bosque, 2001; Corpas, 2001a; Koike, 2001), dado que también determinaría las colocaciones léxicas infrecuentes conocidas como combinaciones libres.

También el enfoque lexicográfico ha cobrado fuerza mediante la figura de Hausmann (1985), quien afirma que en una colocación se da una “relación típica, específica y característica entre dos palabras” (p.118). Su postulado se traduce en una relación de dependencia entre base y colocado, caracterizada porque el

¹⁰ «Modos de significado».

¹¹ Base.

¹² Colocativos o colocados.

primer componente selecciona al segundo. Este aspecto se conoce como “direccionalidad”.

Desde la disciplina de la Lingüística del Texto, Mel'čuk y Zholkovskij (1970) proponen la Teoría Sentido-Texto (TST). Cualquier acto de comunicación lingüística conlleva un contenido (sentido), una representación oral o escrita (texto) y un conjunto de correspondencias entre estos dos elementos. La TST usa las funciones léxicas para establecer un sistema universal, útil para todas las lenguas, que permita formalizar todas las relaciones sintagmáticas y paradigmáticas de cada unidad léxica (Alonso, 1993). Esta teoría tuvo su fruto en 1984 cuando Mel'čuk aplica estos postulados a la creación del primer volumen de un diccionario combinatorio «Explanatory Combinatorial Dictionary»¹³. Logró desarrollar una metodología lexicográfica que describe de forma sistemática las relaciones léxicas establecidas entre las diferentes entradas y que permite al usuario conocer los conceptos y la forma en que se relacionan.

A partir de la década del 70 comienza la aplicación de un enfoque teórico al estudio de las colocaciones. Coseriu (1977) las denomina “solidaridades léxicas” y señala que el significado de una unidad léxica se encuentra definido por la clase semántica a la que pertenece y por su lexema. Descarta la relación establecida entre palabras. Cruse (1986), por su parte, define las colocaciones como “secuencias de unidades léxicas que coocurren habitualmente, no obstante, son completamente transparentes en el sentido de que cada constituyente léxico es también un constituyente semántico” (p. 40). Distingue tres tipos de restricciones colocacionales dependiendo del grado en el que las propiedades semánticas de la base predicen la aparición de un colocado en particular: sistemáticas, semi-sistemáticas e idiosincrásicas.

En el ámbito del español, el concepto de colocación fue introducido por Seco (1978) en una investigación que tenía por objetivo analizar las carencias de los métodos de definición lexicográfica usados en su Diccionario por la Real Academia Española (RAE, 2018).

¹³ «Diccionario de combinaciones explicadas».

La aceptación del término no fue inmediata ni generó reconocimiento unánime en la comunidad científica. Se prefirieron otros conceptos tales como: “solidaridades léxicas”, “lexías complejas”, “combinaciones”, “combinaciones recurrentes”, “conurrencias”, “coocurrencias” o “coapariciones”. (Aguilar-Amat, 1993: 267). Ya en los años 90, diversos hispanistas publicaron sus investigaciones acerca de las colocaciones. Han destacado los aportes de Írsula (1994), Corpas (1996), Koike (2001), Zuluaga (1990) y Alonso (1994).

Otro enfoque que se ha sumado con fuerza en el mundo anglosajón es el de la didáctica, proveniente de la LA. Concibe la colocación como una categoría más amplia que se aplica tanto a combinaciones usuales como a combinaciones probables (Fernández, 2014) con la finalidad de mejorar el proceso de enseñanza y aprendizaje de aprendientes de segundas lenguas.

4.2. Definiciones en torno al concepto de colocación

El Collins COBUILD Dictionary (2006)¹⁴ utiliza la siguiente acepción para el objeto de estudio: “In linguistics, a collocote of a particular word is another word which often occurs with that word”¹⁵. De esta afirmación, se desprende que un criterio fundamental para observar e identificar una colocación es la frecuencia con la que las palabras concurren en un corpus o la aparición de tales unidades. Al respecto, Halliday (1961, 1966) plantea la *probabilidad estadística* de aparición como indicador para determinar la presencia de una colocación.

Una colocación es una estructura que cuenta con una base y un colocado. La base es la palabra principal que constituye el núcleo temático y el colocado, la palabra o frase que la completa y permite su significado y total. Una consideración pertinente es la que realiza Alonso (1994) al señalar que “el término colocación se usa con diferentes sentidos. Algunas veces, se refiere a combinaciones probables

¹⁴ Diccionario Collins COBUILD.

¹⁵ “En lingüística, la colocación de una palabra en particular es otra palabra que suele aparecer con esa palabra”.

o usuales de dos palabras; otras, a combinaciones restringidas en donde un lexema exige la presencia de otro” (p.9). Por su parte, Zuluaga (2002) con el objeto de enriquecer las conceptualizaciones sobre colocaciones señala que las relaciones hipotácticas de estas estructuras son fijadas por el uso lingüístico social, es decir, presentan una relación de dependencia o subordinación en relación a una comunidad determinada de hablantes.

Corpas (1996) entiende por colocación:

las unidades fraseológicas formadas por dos unidades léxicas en relación sintáctica, que no constituyen, por sí mismas, actos de habla ni enunciados; y que, debido a su fijación en la norma, presentan restricciones de combinación establecidas por el uso, generalmente de base semántica (p.76).

Para Koike (2001), la base se refiere a la unidad léxica autónoma semánticamente que selecciona la acepción de la otra unidad de la colocación. Para Hausmann (1979), es el mismo elemento denominado colocador. El colocativo se refiere a la “unidad léxica que se coloca en la periferia de la otra unidad de la colocación” (Koike, 2001, p. 55); en otras palabras, es el elemento que acompaña a la base. Hausmann (1979) lo denomina colocado. La relación establecida se enmarca dentro de una capacidad denominada colocabilidad, definida como “la capacidad que tiene una unidad léxica de combinarse con otra formando así una colocación” (Koike, 2001, p.16). Esta capacidad es gradual y se desarrolla a medida que un aprendiente de LE va adquiriendo léxico de manera individualizada y en bloque.

Existen algunos factores que influyen en la producción y en el uso de las colocaciones. Estos son: 1) Competencia en la segunda lengua (Zhang, 1993; Al-Zahrani, 1998) que se traduce en el dominio de competencia por parte del aprendiente, 2) Influencia de una primera lengua (Granger, 1998a; Alzhrany, 1998; Kaszubski, 2000; Burgschmidt y Perkins, 1985; Kellerman, 1978) que puede afectar de manera total o parcial la producción de las unidades léxicas (Farghal y Obiedat, 1995), 3) Distancia entre la primera lengua y la segunda (Shei, 2000), 4)

Metodología utilizada en la enseñanza de segundas lenguas (Biskup,1992; Felix, 1981; Lightbown,1991), la cual es relevante, ya que influye en el proceso de aprendizaje y adquisición de las colocaciones.

El problema terminológico presente en la fraseología se ha transferido con el paso del tiempo a las colocaciones. No existe un consenso sobre su conceptualización. A continuación, se recogen las definiciones más conocidas en torno a este fenómeno lingüístico.

1. Las colocaciones son combinaciones léxicas libres, formal y semánticamente regulares (Zuluaga, 2002; Muñiz, 2002; García-Page, 2001).

2. Las colocaciones son combinaciones libres de palabras, pero “habitualizadas”, “usualizadas” o “socializadas”, de modo que presentan cierto grado de lexicalización en la norma (Haensch, 1985; Wotjak 1998).

3. Las colocaciones son combinaciones frecuentes en el uso, pero claramente composicionales, y por tanto no lexicalizadas (Corpas, 1996, 2001a; Piera y Varela 1999; Koike, 2001, 2002).

4. Las colocaciones son combinaciones preferentes o habitualizadas que presentan ciertas restricciones combinatorias impuestas por el uso (Mendivil 1991; Aguilar- Amat, 1993; Írsula, 1994; Corpas, 1996).

5. Las colocaciones son un tipo de unidad fraseológica, a medio camino entre las combinaciones libres y las expresiones completamente fraseologizadas; en consecuencia, presentan cierta fijación interna y son parcialmente composicionales (Alonso, 1994, 2002; Blasco, 2002; Koike 2001, 2002; Zuluaga, 2002).

6. Las colocaciones, así como las fórmulas rutinarias y las locuciones, son frases hechas (Baralo, 2007).

Las definiciones anteriores evidencian que las colocaciones son concebidas desde combinaciones libres hasta combinaciones fijadas por norma en una determinada sociedad. Al respecto, Higuera (2004) postula una definición lo más completa considerando algunas anteriores:

un tipo de unidad léxica (Lewis, 1993; Higuera, 1997), una unidad psicológica para los hablantes de una lengua (Benson *et al*, 1986), formada por dos lexemas (Mitchell, 1971) que coocurren con frecuencia, y que expresan una relación típica entre sus componentes (Koike, 2001). Son parcialmente composicionales porque la base¹⁶ mantiene su sentido, pero el colocativo adopta un sentido especial que sólo posee en presencia del otro elemento (Corpas, 1996; Castillo, 1998; Koike, 2001). Ilustran preferencias de combinación y restricciones impuestas por el uso, a nivel de la norma (Corpas, 1996; Zuluaga, 2002), en unos casos debido al conocimiento del mundo y en otros, por razones idiosincrásicas de cada lengua (p.71).

Gairns y Redman (1986) aportan la siguiente definición de colocación: “when two items co-occur or are used together frequently, they are said to collocate” (p. 37)¹⁷. Aclaran que agrupar las unidades léxicas según las propiedades colocacionales resulta difícil, por tanto, aconsejan que se traten conforme vayan surgiendo o de forma aislada. McCarthy (1990) no concibe la colocación como una unidad léxica pluriverbal, es decir, aquella que contiene más de un vocablo, sino más bien como una relación de sentido a nivel paradigmático al igual que la sinonimia, antonimia o hiponimia. Señala que la colocación actúa de principio organizador del léxico de una lengua y que resulta fundamental en el estudio del vocabulario. Al igual que los autores anteriores (Gairns y Redman, 1986; McCarthy, 1990), Nation (1990) no

¹⁶ Palabra clave en la colocación léxica con independencia semántica que selecciona una acepción concreta del colocativo que sólo aparece cuando se combina con esta.

¹⁷ “Cuando dos elementos coexisten o se usan juntos con frecuencia, se dice que se colocan”.

considera la colocación un tipo de unidad léxica, sino que la define como una coaparición de palabras. Este autor aporta a la enseñanza/aprendizaje del léxico en una L2 una de las caracterizaciones más amplias y relevantes sobre lo que significa conocer una palabra. La divide en tres bloques, esto es, la forma, el significado y el uso. En el último bloque es donde precisamente se especifica que conocer una palabra supone saber desde un punto de vista receptivo y productivo qué tipo de elementos usualmente la acompañan.

Más tarde, Lewis (1993, 1997, 2000b) en el enfoque léxico les concede un papel central a las unidades léxicas pluriverbales, entre ellas, la colocación léxica. Destaca la arbitrariedad y la variación en el eje sintagmático y en el paradigmático (Lewis 1993: 93) como principales características.

4.3. Características de las colocaciones

Para Koike (2001), las colocaciones se caracterizan por una serie de rasgos formales y semánticos.

- a) Coocurrencia frecuente: Definida como una de las propiedades más importantes. Corpas (1996) la define como la frecuencia de coaparición de dos unidades léxicas. No obstante; la coocurrencia no es un rasgo exclusivo de las colocaciones.
- b) Restricciones combinatorias: Propiedad impuesta por el uso tradicional. Se manifiesta en la restricción combinatoria en la norma, aunque en el nivel del sistema de lengua los componentes son combinaciones de palabras completamente libres (Corpas, 1996). Por ejemplo, en la colocación “sentir miedo”, la base “miedo” se coloca principalmente con el verbo “sentir”, pero este colocado se combina con distintos sustantivos.

- c) Composicionalidad formal: Permite la flexibilidad de los componentes de la colocación en el nivel morfológico y sintáctico de la lengua. Se puede sustituir un componente de la colocación por su (cuasi-) sinónimo. Por ejemplo, “violiar las normas” por “transgredir las normas” o “vulnerar las normas”. También es posible transformar la colocación en voz pasiva, “condonar la deuda” por “ser condonada la deuda”. Asimismo, las colocaciones admiten modificación del sustantivo con un adjetivo, “dar una explicación” por “dar una explicación satisfactoria”. Se observa la pronominalización en “salir de un apuro” por “salir de él” o la nominalización en uno de los componentes “repicar la campana” por “repique de campanas”. Cabe mencionar que, aunque la colocación permite cierta flexibilidad combinatoria, “las combinaciones libres son más flexibles combinatoria, morfológica y sintácticamente que las colocaciones, por contra, las locuciones son más rígidas” (González, 2006, p.11) y no demuestran tanta flexibilidad.
- d) Vinculación de lexemas: Esta característica da posibilidad de la alternación de la categoría gramatical en los componentes de una colocación, es decir, algunas colocaciones pueden experimentar cambios de categoría gramatical. Corpas (1996) en su manual presenta los ejemplos de los significados léxicos de “*nega*” y “*rotund*” que adquieren diferentes significados¹⁸.
- e) Relación típica entre los componentes: La tipicidad de la relación se expresa en los elementos integrantes. Para Corpas (2001b) “los miembros de las colocaciones reflejan la relación típica, y, por tanto, verdadera que mantienen los colocados en el mundo real” (p. 103). Por ejemplo, forma una colocación “cargar una pistola”, pero no lo es “lavar una pistola” u “olvidar una pistola”, porque el sustantivo pistola sólo puede establecer relación típica en calidad de arma de fuego.

¹⁸ “Negar con rotundidad”, “rotunda negatividad” y “negar rotundamente”.

- f) Precisión semántica de la combinación: Este rasgo semántico quiere decir que la colocación presenta un concepto inequívoco para los hablantes nativos, por lo que desempeña una función fundamental en el acto comunicativo, presenta un concepto particular, indica una “precisión semántica o un concepto inconfundible” (Shirota, 1991, p.23), en otras palabras, esta característica apunta a la distinción entre un concepto y otro.

4.4. Los niveles de las colocaciones

Desde una perspectiva teórica y funcional, existen algunos criterios para determinar una colocación, los que se sitúan en tres niveles: 1) superficial, la distancia que existe entre los colocados y se le conoce como espacio colocacional; 2) textual, la coocurrencia se considera sólo si la colocación aparece en la misma unidad textual y sintáctica; 3) las palabras se encuentran cerca una de otra solo si hay una relación sintáctica directa entre ellas (Evert, 2009).

4.5. Colocaciones léxicas simples y complejas

Para la clasificación de estas combinaciones han sido postuladas dos reconocidas taxonomías. La primera propuesta por Corpas (1996) y la segunda por Koike (2001). Al respecto, Koike (2005) señala que la diferencia entre ambas producciones reside en que:

las colocaciones complejas difieren de las colocaciones simples en su estructura formal y en la distribución cuantitativa. Mientras que en las colocaciones simples las combinaciones sustantivo-verbo son las más representativas en su número, en las complejas las formadas por un verbo y una locución adverbial constituyen el grupo más numeroso (p. 183).

Es decir, el primer grupo está conformado por determinadas clases léxicas, por ejemplo, adjetivo+ sustantivo, a diferencia del segundo grupo, el cual contempla clases léxicas más combinaciones derivadas a partir de estas, por ejemplo, las locuciones. En la Tabla 1, se muestra las combinaciones categoriales y los ejemplos de las colocaciones léxicas simples. Las bases aparecen marcadas en negrita y principalmente corresponden a sustantivos y adjetivos.

Combinación categorial	Ejemplos
1. Sustantivo + adjetivo/ adjetivo+ sustantivo	Discusión acalorada/ santa paciencia
2. Verbo+ sustantivo	Implorar misericordia
3. Verbo+ preposición + sustantivo	Intervenir en un conflicto
4. Sustantivo+preposición+ sustantivo	Banco de peces
5. Adverbio+ adjetivo/ participio	Firmemente convencido
6. Verbo+ adjetivo	Salir mal parado
7. Verbo + adverbio	Cerrar herméticamente

Tabla 1. Clasificación de colocaciones léxicas simples (basada en Corpas 1996 y Koike, 2001).

Las combinaciones categoriales y los ejemplos de las colocaciones léxicas complejas son mostradas en la Tabla 2. Las bases aparecen marcadas en negrita y corresponden a locuciones nominales, sustantivo, verbo y adjetivo.

Combinación categorial	Ejemplos
1. Verbo+ locución nominal	Dar un golpe de estado
2. Locución verbal + sustantivo	Llevar a cabo un proyecto
3. Sustantivo+ locución adjetival	Dinero contante y sonante
4. Verbo+ locución adverbial	Pagar a tocateja
5. Adjetivo+ locución adverbial	Loco de remate

Tabla 2. Clasificación de colocaciones léxicas complejas (basada en Corpas, 1996 y Koike, 2001).

Las colocaciones léxicas simples verbo-nominales se encuentran categorizadas en dos grupos: construcciones con verbos de apoyo (CVA) y construcciones con verbos de significado pleno. También en este sentido Írsula (1994) manifiesta que este tipo es el más numeroso de todas aquellas formadas por la unión de una unidad verbal y un sustantivo. Por su parte Corpas (1996) explica que “este tipo de colocaciones pueden deberse a que los sustantivos presentan mayores posibilidades de modificación” (p. 70) que otras clases léxicas.

4.6. Tipos de colocaciones verbo-nominales

Las colocaciones léxicas con verbos de apoyo pueden ser definidas como aquellas estructuras en las que el colocado cumple una función de auxiliar. Son definidas como un tipo de construcciones verbales en las que el verbo apenas tiene significado léxico (Alonso, 2004). En este mismo sentido, Castillo (2013) las identifica como grupos verbales constituidos por un verbo y un sintagma nominal en los que el contenido proposicional aparece en el sintagma nominal, ya que el verbo tiene, o parece tener, un significado léxico casi nulo; se trata de construcciones como “*dar un paseo*”, “*dar un beso*” en las que el verbo no parece aportar nada más que información gramatical (tiempo, número, persona, aspecto y modo). Blanco (2000) concibe un verbo de esta naturaleza como “un verbo semánticamente vacío que “conjuga” el nombre al que acompaña” (p. 99).

La Nueva Gramática de la Lengua Española (2014) define las CVA como perífrasis verbales de naturaleza semilexicalizada, lo que las sitúa entre dos planos de la lengua. Por una parte, en la sintaxis, y por otra, en el léxico. Por ejemplo, locuciones y colocaciones (Lozano, 2013). Por su parte Alonso (2004) afirma que las CVA son aquellas que sirven de soporte sintáctico al sustantivo que los complementa. En este tipo de colocaciones léxicas verbo-nominales, el verbo desempeña una función deseslematizada, porque es el sustantivo el que expresa el predicado semántico equivalente al significado (Beas, 2009), es decir, el sustantivo, es la base de la estructura. A modo de ejemplo, el autor afirma que hacer es uno de los verbos de apoyo relativamente productivo en las lenguas romances. Desde esta perspectiva, el verbo hacer cumple la función de colocado cuando le sigue una palabra o frase que desempeña función sustantiva. Para De Miguel (2011), este tipo de construcciones:

[...] es el resultado de la interacción de los rasgos subléxicos de un verbo y un sustantivo que constituyen un predicado complejo. Ser verbo de apoyo, desde esta perspectiva, no es otra cosa que ser un verbo con el contenido requerido para combinarse con sustantivos predicativos. [...] en una CVA existe una negociación, un proceso general de concordancia de los rasgos subléxicos de V y N que genera los distintos significados de la construcción en función de la aportación de uno y otro (p.145).

La definición anterior evidencia el grado de atracción entre la base y el colocado, es decir, aquella negociación y co-dependencia entre el sustantivo y el verbo.

En cuanto a sus características, los principales rasgos de una CVA son: a) los sustantivos que constituyen estas estructuras suelen ser eventivos deverbales, y por tanto heredan la estructura argumental del verbo (“dar cumplimiento al reglamento”/ “cumplir el reglamento”); b) la CVA suele ser sinónima de una nominalización con el sustantivo predicativo (“Luis dio un apasionado beso a Marta”/ “el apasionado beso de Luis a Marta”), c) las CVA pueden ser equivalentes a un verbo pleno (“dar cumplimiento”/ “cumplir”; “dar un paseo”/ “pasear”; “hacer

alusión”/ “aludir”; “tomar una decisión”/ “decidir”). Esta última característica es relevante, tanto así que para algunos autores constituye uno de los criterios determinantes para distinguir una CVA de cualquier otro tipo de grupo verbal semilexicalizado (Piera y Valera, 1999) o lexicalizado. Para De Miguel (2008 y 2011) lo característico de estas construcciones no radica en el “vacío léxico” del verbo, sino en que sustantivo y verbo forman una sola unidad predicativa, en la que el sustantivo “concuerta léxicamente” con el verbo; según esta hipótesis el verbo no está “vacío de significado, sino que contiene rasgos mínimos de su definición, no solo aspectuales, que determinan su posibilidad de combinación con ciertos sustantivos” (p. 144) ya sea en corpus lingüísticos orales o escritos.

Para la identificación de un verbo de esta naturaleza, Piera y Varela (1999) proponen que el criterio más claro para identificar qué verbos pueden ser de apoyo consiste en verificar si gracias a ellos se puede construir un sintagma (<verbo de apoyo + sintagma nominal>) equivalente a un verbo sencillo de la lengua, esto quiere decir, utilizar el procedimiento de verbalización en el sustantivo que conforma la colocación.

A diferencia de las colocaciones léxicas simples con verbos de apoyo, los verbos de significado pleno o verbos no copulativos poseen significado y una estructura argumental. No cumplen el rol auxiliar como colocativo, sino que contribuyen a portar significado presentando un contenido semántico determinado. Alonso (1994) considera que los verbos con estas características realizan funciones léxicas, específicamente, funciones verbales fasaes, causativas y realizativas. Desde la perspectiva de Alonso (2004):

a diferencia del verbo de apoyo, que es semánticamente vacío desde un punto de vista sintagmático, el de realización proporciona un sentido específico a la construcción: realizar o llevar a cabo el objetivo inherente a la situación lingüística expresada por el nombre o hacer con aquello

designado por el nombre lo que se espera que se haga (Alonso, 2004: p.114).

Por ejemplo, la colocación “ganar un juego” está constituida por la base “juego”, la cual selecciona semánticamente al colocado “ganar”. Ambos elementos aportan un significado específico, aunque la base contiene el que se considera como elemental.

A la luz de los planteamientos teóricos expuestos hasta este punto, se puede afirmar que los principales fundamentos que explican cómo sucede el aprendizaje o adquisición de las colocaciones léxicas simples verbo-nominales con verbos de apoyo y verbos de significado pleno están basados en el aprendizaje de las unidades léxicas (Izquierdo, 2004), en el enfoque lexicográfico (Hausmann, 1985) y en el enfoque léxico (Lewis, 1993, 1997, 2000b).



5. Lingüística de Corpus

La LC constituye una metodología para la investigación en la adquisición y aprendizaje de las lenguas (Parodi, 2008), con la cual los investigadores recopilan un conjunto de textos en entornos naturales. Las unidades escritas deben poseer sentido y una clara intención comunicativa. Posteriormente, se realiza un análisis cuantitativo y/o cualitativo con la finalidad de evidenciar la problemática objeto de estudio.

Sinclair (1991) sostiene que un corpus es “una colección de textos de ocurrencias de lenguaje natural, escogidos para caracterizar un estado o una variedad de lengua” (p.171). La definición anterior se enriquece en algunos aspectos con la propuesta de Crystal (1991) quien define un corpus como:

una colección de datos lingüísticos, ya sea de textos escritos o de transcripciones de habla grabada, los que pueden ser utilizados

como punto de partida para descripciones lingüísticas o como un medio de verificación de hipótesis acerca de una lengua (p.32).

Un corpus consiste en una colección de partes de una lengua que son seleccionadas y ordenadas de acuerdo a explícitos criterios lingüísticos con el fin de ser empleados como ejemplos de esa lengua. Un corpus debe obedecer a criterios de codificación, estandarización y homogenización para responder a tareas específicas de recuperación (EAGLES, 1996). Para Leech (2002), un corpus puede ofrecer información detallada acerca de una lengua particular, pero es imposible recolectar un corpus que abarque todo un sistema lingüístico. Si ese fuera el caso, sería necesario recolectar todos sus usos. De este modo, se debe siempre tener presente que un corpus es solo una colección finita de un universo infinito. Biber, Conrad y Reppen (1998) consideran que todo análisis lingüístico basado en corpórea debe cumplir los siguientes requisitos:

- a) que sea empírico, es decir, que analice los patrones reales de uso en textos naturales (más que patrones de estructura lingüística),
- b) que utilice una amplia colección de textos naturales como base de análisis,
- c) que utilice el ordenador para tal análisis, usando técnicas automáticas e interactivas, y
- d) que dependa de técnicas analíticas tanto cuantitativas como cualitativas.

Dada su naturaleza, los corpus pueden ser clasificados según distintos tipos. Cruz Piñol (2012) propone las siguientes categorías: según el canal de producción, corpus textuales, orales, mixtos; según el grado de representatividad, corpus textuales, corpus léxicos¹⁹, corpus de referencia; según la especificidad de los textos, corpus generales, genéricos, canónicos, sincrónicos, diacrónicos o históricos, cronológicos o periódicos, especializados; según el porcentaje y distribución de los textos, corpus equilibrados, monitor o abierto y cerrados; según los idiomas incluidos, corpus monolingües, bilingües o multilingües, comparables y

¹⁹ Simple corpus.

paralelos; según el nivel de análisis, corpus simples o brutos²⁰, lematizados, analizados morfológicamente, con información sintáctica superficial, analizados sintácticamente²¹ y etiquetados semánticamente. La nomenclatura anterior constituye una clasificación exhaustiva en el marco de la LC y del ELE, no obstante, no incluye explícitamente los corpus lingüísticos en formato electrónico.

5.1. Corpus de Aprendientes de Lenguas en Formato Electrónico

Una precisión que conviene destacar dentro de la comunidad lingüística es la que sitúa el análisis lingüístico del corpus recopilado en el contexto tecnológico mediante el uso del computador. Se puede señalar que la Lingüística de Corpus tuvo su auge durante la década del 60 y del 70, alentada por las posibilidades que la informática ofrecía para procesar y gestionar conjuntos de textos con un número de palabras cada vez mayor. Una de las consecuencias de lo anterior ha sido el gran avance de la Lingüística Computacional, disciplina de la lingüística que ha contribuido con programas y aplicaciones computacionales en el análisis estadístico de los textos. En conjunto con la Lingüística Computacional, la Lingüística de Corpus ha servido como herramienta de investigación. Ha permitido obtener listados de frecuencia y usos de palabras y detectar estructuras lingüísticas habituales con distintas finalidades según la investigación que se esté llevando a cabo. Tal como lo señala Pravec (2002), en el contexto de la aplicabilidad de la Lingüística de Corpus, “el análisis en la enseñanza de lenguas puede utilizarse para fines específicos y en la investigación sobre la adquisición de una primera lengua y el aprendizaje y adquisición de segundas lenguas o lenguas extranjeras” (p.40). Se presenta como una metodología con gran utilidad para las investigaciones realizadas en el ámbito de ELE.

Por un parte, para el Instituto Cervantes (1997-2019)

los llamados “corpus de aprendices”: contienen textos producidos por personas que están aprendiendo una determinada lengua con diferentes lenguas iniciales o familiares, que denominaremos L1 y distintos grados de

²⁰ Raw corpus.

²¹ Treebanks.

conocimiento de la lengua objeto. Estos corpus permiten que el investigador detecte las fortalezas y debilidades del grupo lingüístico estudiado, por lo tanto, los resultados se consideran una fuente valiosa de datos lingüísticos (Buyse, 2011).

Por otra parte, Granger (2003, 2004) propone que un corpus electrónico de aprendientes de lengua es la colección de datos lingüísticos auténticos (textos orales o escritos) donde se constata el uso de la lengua objeto de estudio (L2 o LE). Su desarrollo ha permitido un avance significativo en los estudios de Adquisición de ASL (Granger, 2012; 2015). Granger (2003) señala que son: 1) computarizados, es decir, se pueden analizar usando distintos programas informáticos que permiten la manipulación rápida de los datos a través de funciones de búsqueda, conteo y clasificación) y 2) extensos, lo que implica que constituyen una base más confiable para describir y modelar la lengua que se está aprendiendo).

Los corpus electrónicos fueron originados en la década de 1990, tras el interés por crear y generar recursos que fueran de utilidad para los aprendientes de lenguas. Como resultado, la consolidación de la LC y la evolución de la tecnología y de las metodologías de investigación permitieron el uso de los corpus informatizados como herramientas de investigación, posibilitando el desarrollo de nuevos tipos de investigaciones tales como los análisis frecuencias de coaparición o de patrones de uso lingüístico por parte de un grupo de aprendientes (Granger, 2004; Hasselgard y Johansson, 2011; Sánchez, 2015). Debido a las posibilidades que ofrece en cuanto a almacenamiento y procesamiento de la lengua, el computador entrega los medios para investigar la lengua de los aprendientes; de esta forma, una vez que los datos del aprendiente se han computarizado pueden tratarse con distintos programas informáticos lingüísticos (Dagneaux *et al*, 1998). Granger (2002) afirma que una de las mayores fortalezas de la metodología de corpus computarizados es su pertinencia para desarrollar análisis cuantitativos; además, gracias al progreso tecnológico ha sido posible recolectar datos de los aprendientes en grandes cantidades, guardarlos en el computador y analizarlos

automática o semiautomáticamente. Kennedy (1998) destaca como ventaja que un corpus almacenado en un computador permite encontrar, clasificar y contabilizar unidades con facilidad, tanto como una base para la descripción lingüística como para abordar problemáticas relacionadas con la lengua. Finalmente, se puede señalar que la naturaleza de estos corpus marcó un antes y un después en el desarrollo y la innovación de las investigaciones dedicadas a la lingüística.

Algunos de los corpus electrónicos de aprendientes son:

1) El Corpus Escrito del Español como L2 (CEDEL2) de aprendices de inglés como L1. Este corpus considera producciones textuales en todos los niveles de proficiencia y se recolectó con fines contrastivos entre el español y el inglés (Lozano y Mendikoetxea, 2013). CEDEL2 está conformado por alrededor de 750.000 palabras en formato electrónico y se ha recolectado de forma sistemática vía online con estudiantes y profesores voluntarios. El corpus está disponible para ser utilizado como fuente de datos para los investigadores de ELE y con abundancia de ejemplos para la enseñanza de ELE (Ferreira y Elejalde, 2017). Los textos cuentan con diferentes variables metodológicas: edad, nivel de competencia lingüística, edad de exposición al español, años estudiando español, duración de la estancia para aprender esta lengua, entre otras.

2) El Corpus de Aprendices de Español como Lengua Extranjera (CAES) (Rojo y Palacios, 2015). Corresponde a un conjunto de textos escritos, producidos entre los años 2011 y 2013 por estudiantes de español como LE de todos los niveles de español (A2-B1- B2-C1). Este corpus ha sido elaborado en colaboración con investigadores y docentes de la Universidad de Santiago de Compostela y del Instituto de Cervantes, España. Los textos han sido generados por aprendientes de diferentes lenguas maternas (árabe, chino, mandarín, francés, inglés, portugués y ruso), provenientes de distintas universidades del mundo. El total de textos recolectados es de 3878 provenientes de 1423 estudiantes que escribieron dos o tres textos cada uno. El objetivo del proyecto es permitir el uso de los datos para la investigación en la adquisición de ELE y está disponible para la consulta en línea (Ferreira y Elejalde, 2017).

3) El Corpus de Interlengua Oral de italiano y español como LE (CORINÉI) comprende textos orales (grabación de las interacciones vía Skype, transcripción y almacenamiento de las conversaciones) a partir del desarrollo del proyecto colaborativo Teletándem (González, 2012). Dicho proyecto está orientado a facilitar el aprendizaje colaborativo entre hablantes nativos y no nativos de español e italiano en la enseñanza aprendizaje de lenguas para la traducción. Los resultados han permitido investigar las estructuras oracionales de las conversaciones generadas en este tipo de ambiente de aprendizaje y, de este modo, contribuir al mejoramiento de las habilidades orales en ambas lenguas (Elejalde, 2017).

4) El Corpus de Suecos Aprendices de ELE (SAELE) (Pino, 2009) está conformado por una colección digital de textos argumentativos escritos por aprendices de ELE provenientes de Suecia. Este corpus ha sido recopilado en dos universidades suecas, durante el 2008 al 2009. Los niveles de proficiencia de los aprendices son A2 y B1. En su primera etapa fueron recolectados 135 textos, que fueron escritos por 45 estudiantes de dos universidades suecas. La exploración de este corpus ha permitido obtener las frecuencias de usos en estructuras gramaticales o de coherencia textual, presentes en la interlengua de aprendientes suecos. El estudio realizado por Pino (2012) se centró en el análisis de los conectores “porque, por eso y entonces” en la interacción con diferentes tipos de palabras (Elejalde, 2017).

5.1.1. Corpus CAELE

El Corpus de Aprendientes de Español como Lengua Extranjera CAELE (Ferreira, 2014-2018) ha sido generado en el marco de los proyectos de investigación Fondecyt N° 1140651 (Ferreira, 2014-2018) y Fondecyt N° 1180974 “Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua” (Ferreira, 2014-2018) con el propósito principal de estudiar diferentes fenómenos que atañen a la conformación

de la interlengua de los aprendientes de ELE de acuerdo con la producción más auténtica y real de muestras escritas (Ferreira y Elejalde, 2020). Para el análisis de los textos se utiliza el formato digital txt., el cual es almacenado en el *software Sketch Engine* y *UAM Corpus Tool* en el contexto de la enseñanza del ELE. Esta colección de textos se incrementa año a año. Actualmente se distribuye entre los años 2014 y 2019. Cuenta con 1217 textos, escritos por 201 aprendientes de nivel A2 y B1 en un rango de edad que fluctúa entre los 18 y 26 años. Los textos consideran distintas lenguas maternas, entre las más representativas destaca: francés, 384 textos; inglés, 325 textos y alemán, 312 textos.

Los estudiantes provienen de los programas de intercambio que tiene la Universidad de Concepción (Chile) en el marco de la movilidad estudiantil (Ferreira y Elejalde, 2020). Su nivel de competencia lingüística fue determinado a través de la Prueba Multinivel con Fines Específicos Académicos (Ferreira, 2016), instrumento alineado con los niveles del Marco Común Europeo de Referencia (Consejo de Europa, 2001).



5.1.2. Investigación a partir del Corpus CAELE

A partir de la recolección del corpus CAELE, se han investigado variadas problemáticas lingüísticas. Destacan los siguientes estudios, los cuales sustentan metodológicamente esta tesis.

- **Propuesta de una taxonomía etiológica para etiquetar errores de interlengua en el contexto de un corpus escrito de aprendientes de ELE (Ferreira y Elejalde, 2020)**

El estudio llevado a cabo por Ferreira y Elejalde tiene como objetivo general proponer una taxonomía para la anotación de los errores etiológicos sistemáticos causados por transferencia lingüística de la L1. Para ello, se abordan los principales fundamentos teóricos en materia de tipología de errores (Corder, 1967;

Alexopoulou, 2006, 2007; De Alba Quiñones, 2009) y taxonomías con criterio etiológico (Alexopoulou, 2005, 2006; De Alba Quiñones, 2009; Ferreira, 2014, Ferreira, Elejalde y Vine 2014, Elejalde y Ferreira, 2016) y la descripción de la taxonomía propuesta. Dicha propuesta se concibe como una evolución taxonómica de estudios previos y del proyecto Fondecyt N° 1180974 (Ferreira, 2018-2021). El criterio etiológico permite determinar la fuente del error y entregar una explicación de cuál es su naturaleza. Considera dos categorías: errores intralingüísticos y errores intralingüísticos (Ferreira y Elejalde, 2020).

- **Colocaciones gramaticales verbo + preposición en ELE: aspectos de naturalidad en aprendientes anglófonos de nivel A2 y B1 (Oportus y Ferreira, 2019)**

Este estudio, cuyo diseño es descriptivo y transversal, con un análisis mixto de datos, tiene como propósito examinar el uso de las colocaciones gramaticales v+ preposición (ColGram v+prep) en la escritura de estudiantes anglófonos de ELE de nivel A2 y B1, identificando similitudes y diferencias en la interlengua colectiva de estos participantes en relación con la producción hablantes nativos (HN) de español. En cuanto a la metodología, el estudio se sustenta en el enfoque de investigación del Análisis Contrastivo de la Interlengua (Granger, 2015) y en los procedimientos metodológicos del Análisis de Corpus Electrónicos de Aprendientes desarrollados en los proyectos Fondecyt N°1140651 (Ferreira, 2014-2018) y Fondecyt N°1180974 (Ferreira, 2018-2021). Basados en los datos de frecuencia, se concluye que los aprendientes muestran rasgos de naturalidad restringida en el uso de las unidades estudiadas, en términos que el nivel A2 produce un menor número y menor variedad de unidades que los HN, mientras que el nivel B1, aunque alcanza una frecuencia similar, presenta también un índice de variedad inferior a los HN, tal como el nivel A2 (Oportus y Ferreira, 2019).

- **Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE (Ferreira y Elejalde, 2017)**

Este estudio con un diseño descriptivo y longitudinal tiene como objetivo principal determinar los errores más frecuentes y recurrentes en el CAELE. Para ello se utiliza la metodología de investigación del Análisis de Errores Asistido por Computador. La identificación, clasificación y anotación de los errores se procedió con los criterios y taxonomía de errores del modelo metodológico desarrollado por Ferreira (2011, 2014, 2018). Los hallazgos de esta investigación indican que los errores más frecuentes y recurrentes del CAELE son: la falsa selección de género gramatical y la omisión de tilde ortográfica. En cuanto a los niveles de competencia lingüística A2 y B1, el error más recurrente corresponde a la omisión de tilde ortográfica en las palabras esdrújulas. A modo de conclusión, los resultados presentados sugieren una tendencia importante en el orden en que se deben tratar los errores según el nivel de competencia en el proceso de corrección por escrito. Finalmente, este estudio constituye un aporte valioso en el ámbito de los corpus de textos escritos en formato electrónico de ELE (Ferreira y Elejalde, 2017).

Finalmente, tal como lo plantea Elejalde (2017), el incremento en el interés por este tipo de investigaciones en corpus releva la necesidad de identificar las problemáticas o fenómenos que atañen al uso de la lengua meta en diversos contextos comunicativos. Por consiguiente, los corpus de aprendientes son especialmente útiles a la hora de investigar el proceso de aprendizaje, constituyéndose en una fuente empírica de datos para líneas de investigación como el CEA²² y los estudios en análisis de interlengua Cruz Piñol, 2012; Ferreira y Elejalde, 2017).

²² Análisis de Errores Asistido por Computador.

CAPÍTULO II. METODOLOGÍA DE LA INVESTIGACIÓN

En este apartado se detalla la metodología que permitió llevar a cabo los estudios sobre colocaciones léxicas simples verbo-nominales. La investigación está sustentada principalmente en la metodología de Lingüística de Corpus y en los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico en cuanto a la recopilación del corpus, a su anotación lingüística y al procesamiento de los datos a través de la herramienta computacional *UAM Corpus Tool* versión 3.2. Asimismo, la base metodológica de esta investigación se enmarca en los proyectos Fondecyt N°1140651 (Ferreira, 2014-2018) y Fondecyt N°1180974 (Ferreira, 2018-2021).

6. Pregunta de investigación

La pregunta que condujo el desarrollo de la investigación se enuncia a continuación:

¿Varían las colocaciones léxicas simples con verbos de apoyo y con verbos de significado pleno (colocaciones verbo-nominales) producidas por los aprendientes de niveles A2 y B1 del ELE?

6.1. Hipótesis

En esta tesis se postulan las siguientes hipótesis de trabajo:

- H1: La producción de colocaciones léxicas simples con verbos de apoyo es menor que la producción con verbos de significado pleno.
- H2: La producción de colocaciones léxicas simples verbo-nominales es mayor en el nivel de competencia lingüística B1 que en el nivel A2.

Las hipótesis nulas postuladas son:

- Ho1: La producción de colocaciones léxicas simples con verbos de apoyo no es menor que la producción con verbos de significado pleno.
- Ho2: La producción de colocaciones léxicas simples verbo-nominales no es mayor en el nivel de competencia lingüística B1 que en el nivel A2.

En la Tabla 3 está contenida tanto la definición como la operacionalización de las variables presentes en las hipótesis.

Variab les	Variable nominal	Variable operacional
Colocaciones léxicas simples verbo-nominales	Tipo de combinación frecuente compuesta por verbo transitivo más un sustantivo.	Se mide a través de la frecuencia mínima de tres de estas combinaciones en el sub-corpus CAELE.
Colocaciones léxicas simples verbo-nominales con verbos de apoyo	Tipo de combinación frecuente compuesta por verbo transitivo y auxiliar más un sustantivo.	Se mide a través de la frecuencia mínima de tres de estas combinaciones en el sub-corpus CAELE.
Colocaciones léxicas simples verbo-nominales con verbos con significado pleno	Tipo de combinación frecuente compuesta por verbo transitivo y que porta el contenido semántico más un sustantivo.	Se mide a través de la frecuencia mínima de tres de estas combinaciones en el sub-corpus CAELE.
Nivel de competencia A2	Nivel de competencia inicial para el aprendizaje del ELE.	Se determina a través de la aplicación de la Prueba Multinivel con Fines Específicos Académicos (Ferreira, 2016)
Nivel de competencia B1	Nivel de competencia intermedio para el aprendizaje del ELE.	Se determina a través de la aplicación de la Prueba Multinivel con Fines Específicos Académicos (Ferreira, 2016)

Tabla 3. Definición y operacionalización de las variables. Elaboración propia.

6.2. Objetivos de investigación

6.2.1. Objetivos generales

Los objetivos generales planteados en esta tesis doctoral son:

- Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE.
- Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1.

6.2.2. Objetivos específicos

Los siguientes objetivos específicos contribuyen a la consecución del primer objetivo general.

- Identificar las colocaciones léxicas simples verbo-nominales de competencia del ELE A2 y B1 utilizando los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico.
- Clasificar las colocaciones léxicas simples verbo-nominales en construcciones con verbos de apoyo y en construcciones con significado pleno.
- Delimitar las frecuencias de bases y colocados de las estructuras en estudio.

Los siguientes objetivos específicos que contribuyen a la consecución del segundo objetivo general son:

- Comparar las colocaciones léxicas simples verbo-nominales con verbos de apoyo de acuerdo con los niveles del ELE A2 y B1.
- Comparar las colocaciones léxicas simples verbo-nominales con verbos plenos de acuerdo con los niveles del ELE A2 y B1.
- Comparar las frecuencias de bases y colocados de las estructuras en estudio de acuerdo con los niveles del ELE A2 y B1.

A continuación, se describe cada estudio. Se incluye el tipo de investigación y la metodología utilizando los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico.

6.3. Estudio 1: Frecuencia colocacional en el sub-corpus CAELE

En relación con el primer objetivo general de esta tesis, se llevó a cabo un estudio descriptivo transversal con un enfoque de análisis de datos mixtos. Para Hernández, Fernández y Baptista (2014), “la meta del investigador consiste en describir fenómenos, situaciones, contextos y sucesos; esto es detallar cómo son y se manifiestan” (p.92). Los estudios de esta naturaleza especifican las propiedades y características de cualquier objeto o fenómeno analizado (Hernández, Fernández y Baptista, 2014). Se consideran para su desarrollo las siguientes variables: la frecuencia de las colocaciones léxicas simples verbo-nominales, la frecuencia de las colocaciones léxicas simples con verbos de apoyo, la frecuencia de las colocaciones léxicas simples con verbos de significado pleno y los niveles de competencia lingüística (A2 y B1).

Para la realización de este estudio, se ejecutó la prueba estadística *t-test*²³, para muestras pareadas en relación a los dos tipos de colocaciones léxicas simples verbo-nominales, es decir, colocaciones conformadas con verbos de apoyo y

²³ La prueba se llevó a cabo en la interfaz de Quick Cals, herramienta disponible en la página web: <https://www.graphpad.com/quickcalcs/ttest1/>

colocaciones conformadas con verbos de significado pleno. La meta apunta a determinar si existía una diferencia significativa entre estos dos tipos de colocaciones. En este sentido, la aplicación de la prueba responde la H1: La producción de colocaciones léxicas simples con verbos de apoyo es menor que la producción con verbos de significado pleno.

En este contexto, los procedimientos para confirmar y validar la H1 corresponden a: a) identificación manual de colocaciones léxicas simples verbo-nominales, b) clasificación de colocaciones léxicas simples verbo-nominales en colocaciones con verbos de apoyo o colocaciones con verbos de significado pleno, c) obtención y descripción de resultados preliminares, d) comparación de los dos tipos de colocaciones mediante la prueba *t-test*, e) obtención y descripción de resultados a partir la prueba estadística y f) comparación de los resultados a la luz de la literatura científica en el ámbito de colocaciones léxicas simples verbo-nominales producidas en LM y ELE.



6.3.1. Procedimientos metodológicos

- Delimitación de un sub- corpus con muestras de hablantes del ELE a partir del corpus CAELE.
- Análisis de las colocaciones léxicas verbo-nominales utilizando los procedimientos del *software UAM Corpus Tool*.
- Descripción de las frecuencias colocacionales en los niveles de competencia lingüística A2 y B1.
- Validación del análisis realizado en el estudio de frecuencia colocacional por parte de otros investigadores.

6.3.2. Sub-corpus del corpus CAELE

Para efectos de la investigación, se seleccionaron los textos de los aprendientes con las lenguas maternas con mayor representatividad, es decir, alemán, inglés y francés. Estas permitieron la conformación del sub-corpus. Su tamaño ascendió a 341 textos producidos entre el año 2014 y el año 2016. Del total de textos, 60 fueron recopilados durante el año 2014, 203 el año 2015 y 78 en el 2016. Respecto de la distribución por nivel de competencia, 188 de los textos correspondieron al nivel A2, y 153, al nivel B1 (ver Tabla 4).

Año de recopilación	N° textos		Total
	A2	B1	
2014	23	37	60
2015	111	92	203
2016	54	24	78
Total	188	153	341

Tabla 4. Distribución del sub-corpus en los años de recopilación. Elaboración propia.

La Tabla 5 reúne a modo de presentación general las características, la recolección del sub-corpus y las variables que intervinieron según los planteamientos metodológicos propuestos por Tono (2003) y adaptados por Ferreira *et al* (2014). Cada uno de estos elementos constitutivos es descrito exhaustivamente.

Rasgos de los textos		Recolección de los datos		Variables del aprendiente	
Modo	Escrito	Tipo y modo de recolección	Concurrente Digital	Internas:	24 años
Género	Texto informativo	Elicitación	Por instrucción	L1	Alemán, francés e inglés.
Estilo	Narrativo/ Descriptivo	Uso de referencias	Sin referencia	L2	ELE
Tópico	Específico	Limitación de tiempo	30 minutos	Nivel de competencia	A2 B1

Tabla 5. Diseño y consideraciones para la construcción de un corpus de aprendientes (adaptada y traducida por Ferreira, Elejalde y Vine (2014), de Tono (2003).



6.3.3. Rasgos de los textos

Este corpus se caracterizó por la escritura de textos con un estilo narrativo/descriptivo que abordó temáticas específicas acordes con el nivel de competencia lingüística de los aprendientes. Estas se visualizan en la Tabla 6. Los tópicos más productivos en el nivel A2 fueron: “artesanía chilena”, “literatura país de origen”, “música folclórica país de origen” y “pinturas costumbristas”. En el nivel B1, la temática, cuya selección reunió mayor cantidad de textos fueron: “artesanía chilena”, “literatura país de origen”, “música folclórica país de origen” y “música folclórica chilena”.

N°	Temas	Distribución de tareas de escritura		
		A2	B1	Total
1	Artesanía país de origen	6	8	14
2	Artesanía chilena	22	25	47
3	Literatura país de origen	21	21	42
4	Literatura latinoamericana	4	7	11
5	Música folclórica país de origen	20	18	38
6	Música folclórica chilena	16	18	34
7	Pinturas costumbristas	18	14	32
8	Música popular chilena	6	4	10
9	Música juvenil	16	13	29
10	Leyendas país de origen	4	3	7
11	Turismo país de origen	6	4	10
12	Lugares turísticos de Latinoamérica	9	3	12
13	Lugares turísticos de Chile	12	4	16
14	Pueblos originarios	12	10	22
15	Música latinoamericana	4	0	4
16	Artesanía latinoamericana	3	1	4
17	Relaciones amorosas	3	0	3
18	Políticamente correcto	3	0	3
19	Fiestas	3	0	3
-	Total	188	153	341

Tabla 6. Distribución de tareas de escritura por tema textual. Elaboración propia.

La extensión es un rasgo importante que se añadió para ampliar las características del sub-corpus. Se les solicitó a los estudiantes que sus redacciones debían cumplir como requerimiento con un número mínimo de palabras, acorde al nivel de competencia lingüística (150 para A2 y 200 para B1). La Tabla 7 considera el número de palabras, su extensión mínima y máxima, y su promedio por texto.

Rasgos	A2	B1	Total sub-corpus
Nº total de palabras	35030	35605	70635
Tokens	38989	39520	78509
Extensión mínima palabras	155	210	-
Extensión máxima palabras	422	536	-
Promedio palabras por texto	186	232	-

Tabla 7. Número de palabras y *tokens* del sub-corpus de aprendientes. Elaboración propia.

En la Tabla 8 se muestra la distribución de los textos según la lengua materna de los aprendices. Se puede observar que tanto en A2 como en B1 la mayoría de los textos correspondieron a aprendientes cuya L1 es alemán, seguidos del inglés, el francés y el portugués.

Lengua	Distribución de textos según LM		Total
	A2	B1	
Inglés	68	40	108
Alemán	83	67	150
Francés	25	32	57
Portugués	12	14	26
Total	188	153	341

Tabla 8. Distribución de los textos según lengua materna. Elaboración propia.

6.3.4. Recolección de los datos

El sub-corpus lingüístico fue recopilado en el Laboratorio de Aplicaciones Tecnológicas de Lingüística Aplicada (LATLA) en el contexto del Proyecto Fondecyt N° 1180974 (Ferreira-2018-2021). Los aprendientes, a lo largo del curso, debieron producir distintas tareas de escritura en un modo digital con temáticas alusivas a Chile y a su país de origen. Las producciones fueron escritas en un tiempo aproximado de 30 minutos.

Los aprendientes produjeron sus redacciones en distintos momentos del semestre académico. Cada tarea de escritura fue escrita sin ningún uso de diccionario ni traductor y con instrucciones explícitas proporcionadas por el docente a cargo.

6.3.5. Descripción de los aprendientes de ELE

Los aprendientes que generaron el sub-corpus fueron 62 estudiantes universitarios extranjeros de intercambio (*ver Tabla 9*) en los periodos contemplados desde el año 2014 al 2016 y cuyo nivel de competencia correspondió a A2 y B1 según la Prueba de Multinivel con Fines Específicos Académicos (Ferreira 2016), instrumento alineado con los niveles del Marco Común Europeo de Referencia (Consejo de Europa, 2001) para las Lenguas.

Datos de la muestra	A2		B1		Total sub-muestra
	Hombres	Mujeres	Hombres	Mujeres	
Número participantes	15	14	19	14	62

Tabla 9. Género de los aprendientes de ELE. Elaboración propia.

En el nivel A2, el 52% de los participantes tenía género masculino, mientras que el 48% restante tenía género femenino. En cuanto al nivel B1, el 58% tenía género masculino y el 42%, género femenino. El promedio de edad fue de 24 años aproximadamente. Los sujetos provenían de distintas carreras, tales como

ingeniería, matemáticas e industria administración y astrofísica, áreas más relacionadas con las ciencias exactas y el uso del lenguaje matemático. También provenían del área de las humanidades y ciencias sociales (Ferreira y Elejalde, 2019).

Los aprendientes se caracterizaron por presentar un estilo de aprendizaje activo y participativo. Sus motivaciones radicaron en desarrollar la competencia comunicativa del español como L2 en un contexto auténtico de aprendizaje en las dimensiones de comprensión auditiva, comprensión de lectura, léxico, gramática, producción escrita y producción oral. Cada participante escribió entre 4 y 7 textos.

Tal como se indica en la Tabla 10, aproximadamente la mitad de los aprendientes del ELE (47%) produjo sus textos durante el año 2015, seguidos por un 37%, correspondiente al año 2016. Finalmente, un 16 % produjo sus tareas de escritura en el primer año de recopilación.

Año de recopilación	N° de aprendientes		Total
	A2	B1	
2014	6	4	10
2015	13	16	29
2016	14	9	23
Total	33	29	62

Tabla 10. Distribución de los aprendientes durante los años de recopilación del sub-corpus. Elaboración propia.

6.4. Estudio 2: Comparación de la producción colocacional

En relación al segundo objetivo general planteado en esta tesis, se desarrolla un estudio comparativo sobre la producción y los usos de las colocaciones léxicas simples verbo-nominales en los dos niveles de competencia lingüística. En este sentido, un estudio de esta naturaleza puede ser definido como aquel que

selecciona procedimientos de búsqueda sistemática para establecer similitudes léxicas en una problemática lingüística determinada.

Para la realización de este estudio, se ejecutó la prueba estadística para muestras independientes *t-test*, cuya meta es observar si existía una diferencia significativa en dos grupos, en este caso, en los grupos separados por los niveles de competencia del ELE (A2 y B1). Los resultados de su aplicación, valores considerados como significativos o no significativos, permiten evaluar la confirmación o rechazo de la segunda hipótesis postulada en esta investigación, es decir, H2: La producción de colocaciones léxicas simples verbo-nominales es mayor en el nivel de competencia lingüística B1 que en el nivel A2. Las variables implicadas en este estudio corresponden a la frecuencia de las colocaciones verbo-nominales, la frecuencia de los verbos de apoyo, la frecuencia de los verbos de significado pleno, niveles de competencia (A2 y B1) y la competencia lingüística de hablantes nativos de español.

Como parte de los procedimientos que permiten observar los usos de las colocaciones estudiadas en lengua materna, también se ejecutó la prueba ANOVA²⁴, para medir la diferencia de la producción colocacional entre los 2 grupos estudiados (aprendientes de niveles A2 y B1) con la de un grupo de hablantes nativos.

En cuanto a los procedimientos para confirmar y validar la H2, estos son a) identificación manual de colocaciones léxicas simples verbo-nominales, b) clasificación de colocaciones léxicas simples verbo-nominales en colocaciones con verbos de apoyo o colocaciones con verbos de significado pleno, c) clasificación de los componentes (bases y colocados), d) clasificación de los tipos de colocaciones y sus componentes según niveles A2 y B1 del ELE, e) obtención y descripción de resultados preliminares, f) comparación de los dos tipos de colocaciones según niveles A2 y B1 del ELE mediante la prueba *t-test*, g) comparación de los resultados a la luz de la prueba estadística ANOVA para medir

²⁴ Esta medida estadística se efectuó en la interfaz Good Calculator, herramienta disponible en la página web: <https://goodcalculators.com/one-way-anova-calculator/>

la diferencia de la producción colocacional entre los 3 grupos (aprendientes de nivel A2, aprendientes de nivel B1 y hablantes nativos), h) obtención y descripción de resultados a partir la prueba estadística e i) comparación de los resultados a la luz de la literatura científica en el ámbito de colocaciones léxicas simples verbo-nominales producidas en LM y ELE.

6.4.1. Procedimientos metodológicos

- Delimitación de un sub- corpus con muestras de hablantes de nativos de español.
- Análisis las colocaciones léxicas verbo-nominales utilizando los procedimientos del *software UAM Corpus Tool*.
- Comparación entre niveles de competencia lingüística de las colocaciones léxicas verbo-nominales del sub-corpus CAELE.
- Comparación de las colocaciones léxicas obtenidas en el sub-corpus del Corpus CAELE con las colocaciones del sub-corpus de español.
- Validación del análisis realizado en el estudio comparativo por parte de otros investigadores.

6.4.2. Corpus de hablantes nativos de español

El corpus de nativos es un corpus que fue recolectado en el año académico 2018. Fue desarrollado en el marco de la tesis doctoral, realizada por Kloss (2018). Está compuesto por 164 tareas de escritura escritas en español por 35 estudiantes, cuya edad fluctuaba entre los 18 y 22 años. Cursaban primer año de la carrera de periodismo en una universidad chilena.

6.4.3. Sub-corpus de español

Con la finalidad de generar un corpus comparable, fueron seleccionados 83 textos que contaban con una cantidad de palabras similar a las tareas de cada escritura de cada nivel. Sus características resumidas se explicitan en la Tabla 11.

Rasgos de los textos		Recolección de los datos		Variables del aprendiente	
Modo	Escrito	Tipo y modo de recolección	Concurrente Digital	Internas:	21 años
Género	Ensayo, columna de opinión y motivacional	Elicitación	Por instrucción	L1	Español
Estilo	Argumentativo/Expresivo	Uso de referencias	Sin referencia	L2	No aplica
Tópico	equidad de género y Libre	Limitación de tiempo	30 minutos	Nivel de competencia	No aplica

Tabla 11. Diseño y consideraciones para la construcción de un corpus de aprendientes (adaptada y traducida por Ferreira, Elejalde y Vine (2014), de Tono (2003).

6.4.4. Rasgos de los textos

Los textos fueron producidos en un modo escrito y abarcaron 2 géneros discursivos, 2 estilos y 3 tópicos (equidad de género y 2 temáticas libres, una referida a una opinión y la otra referida a un texto motivacional). Estas se enmarcaron en el programa de la asignatura Producción de textos.

Al igual que en el sub-corpus de ELE, se añadió como rasgo textual información numérica sobre la extensión, la cual se encuentra distribuida en la Tabla 12

Rasgos	Total corpus
Nº total de palabras	34942
Tokens	38751
Extensión mínima palabras	223
Extensión máxima palabras	1379
Promedio palabras por texto	420

Tabla 12. Número de palabras y *tokens* del sub-corpus de nativos del español. Elaboración propia.

6.4.5. Recolección del sub-corpus

La recolección de los textos se llevó a cabo en un laboratorio de computación en el contexto de la asignatura que cursaban los estudiantes, la cual se impartió bajo modalidad presencial e incluía actividades de escritura redactadas en el aula virtual. Las actividades se realizaron en el marco dado por los objetivos y contenidos del programa de asignatura y comprendían tareas calificadas, así como de diagnóstico y proceso, sin calificación. Los estudiantes escribieron sus textos directamente en el computador en periodos de tiempo aproximados de 30 minutos, bajo la supervisión de la docente responsable.

6.4.6. Descripción de los hablantes nativos de español

Los hablantes nativos de español fueron 35 estudiantes. Un 63% correspondió a hombres, mientras que un 37% correspondió a mujeres (*ver Tabla 13*). El promedio de edad fue de 21 años aproximadamente. Cada participante escribió un total de 3 textos.

Datos de la muestra	Hombres	Mujeres	Total muestra
Número participantes	22	13	35

Tabla 13. Género de los hablantes nativos. Elaboración propia.

6.5. Procesamiento de los sub-corpus lingüísticos

Los sub- corpus fueron procesados en el programa *UAM Corpus Tool* versión 3.2. Antes de subir los archivos al *software*, se requirió que todos los textos estuvieran en formato bloc de notas o txt. y que fueran sometidos a un proceso de limpieza que consistió en aplicar justificado de línea, eliminar tildación proveniente de otra lengua y, finalmente, almacenar el archivo en formato UTF-8.

A cada archivo se le asignó una serie. Para el sub-corpus de ELE, estuvo constituida por el año, número de sujeto, el nivel de competencia y el número de tarea, por ejemplo: 2015S18B1T1. Para el sub-corpus de nativos, la serie estuvo constituida por el año, el número de sujeto, el número de tarea y el corpus, por ejemplo: 2018S20T4CHN.

6.5.1. Software UAM Corpus Tool 3.2

El *software UAM Corpus Tool*²⁵ es una herramienta desarrollada por O' Donnell (2008). Permite el procesamiento de datos a través de la creación de una o más taxonomías, el etiquetado del documento con categorías específicas (año, nivel de competencia lingüística, lengua materna, etc.) y el etiquetado, en distintos niveles

²⁵ Este software fue desarrollado en el contexto del Ministerio de Educación y Ciencia de España bajo el número de licencia HUM2005-01728/FILO (The WOSLAC Project). Para mayor información consultar www.wagsoft.com/Corput/Tool.

de análisis (morfológico, léxico, semántico, sintáctico, pragmático), de una categoría léxica en particular, una frase, oración o párrafo.

Los resultados son automáticos, se expresan en frecuencias y se encuentran accesibles por medio de tres formatos de matrices: (1) *describe a data set*²⁶; (2) *compare several data sets*²⁷, y (3) *describe each file*²⁸. Al etiquetar, se genera una copia del texto analizado. De esta forma, se conserva el texto original, el cual se puede editar sin problemas.

6.6. Taxonomía para la anotación de colocaciones léxicas

La taxonomía implementada forma parte de la Taxonomía para clasificación de errores desarrollada en los Proyectos Fondecyt N°1110812 (Ferreira, 2011-2014) y N° 1180974 (Ferreira, 2018-2021).

En la Figura 1 se visualiza parte de la taxonomía organizada jerárquicamente en el *software* utilizado. Esta se enfoca en definir una colocación v+s y luego delimitar si está formada por un verbo de significado pleno o por un verbo de apoyo.

²⁶ Descripción general del corpus.

²⁷ Descripción del corpus comparando dos o más variables.

²⁸ Descripción de cada archivo digital del corpus.

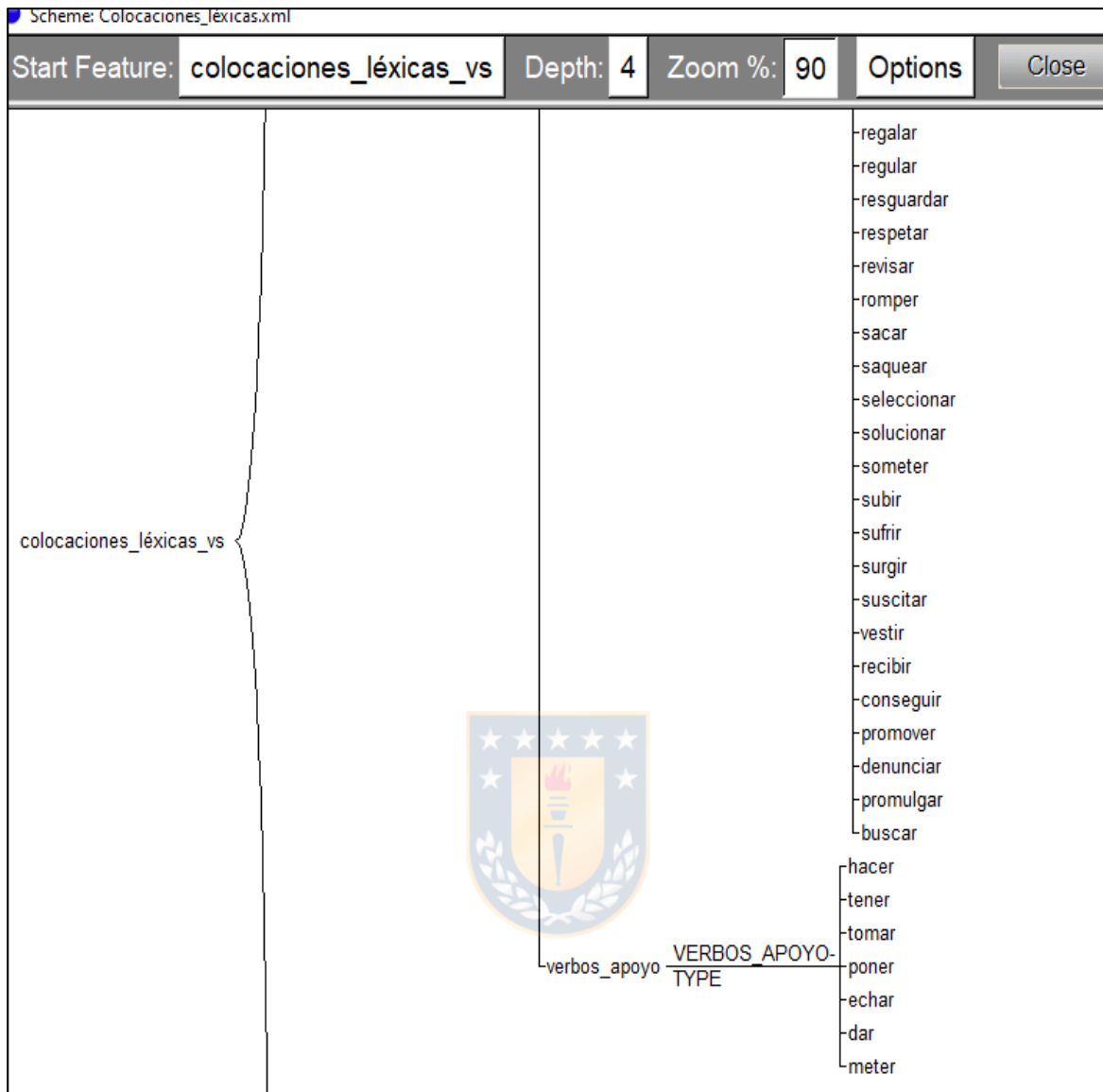


Figura 1. Taxonomía estudio de frecuencia colocacional.

En tanto en la Figura 2, se muestra la interfaz de aplicación de la taxonomía en un texto, perteneciente a un aprendiente de ELE (20014S11A2POSIN) para la anotación de colocaciones léxicas

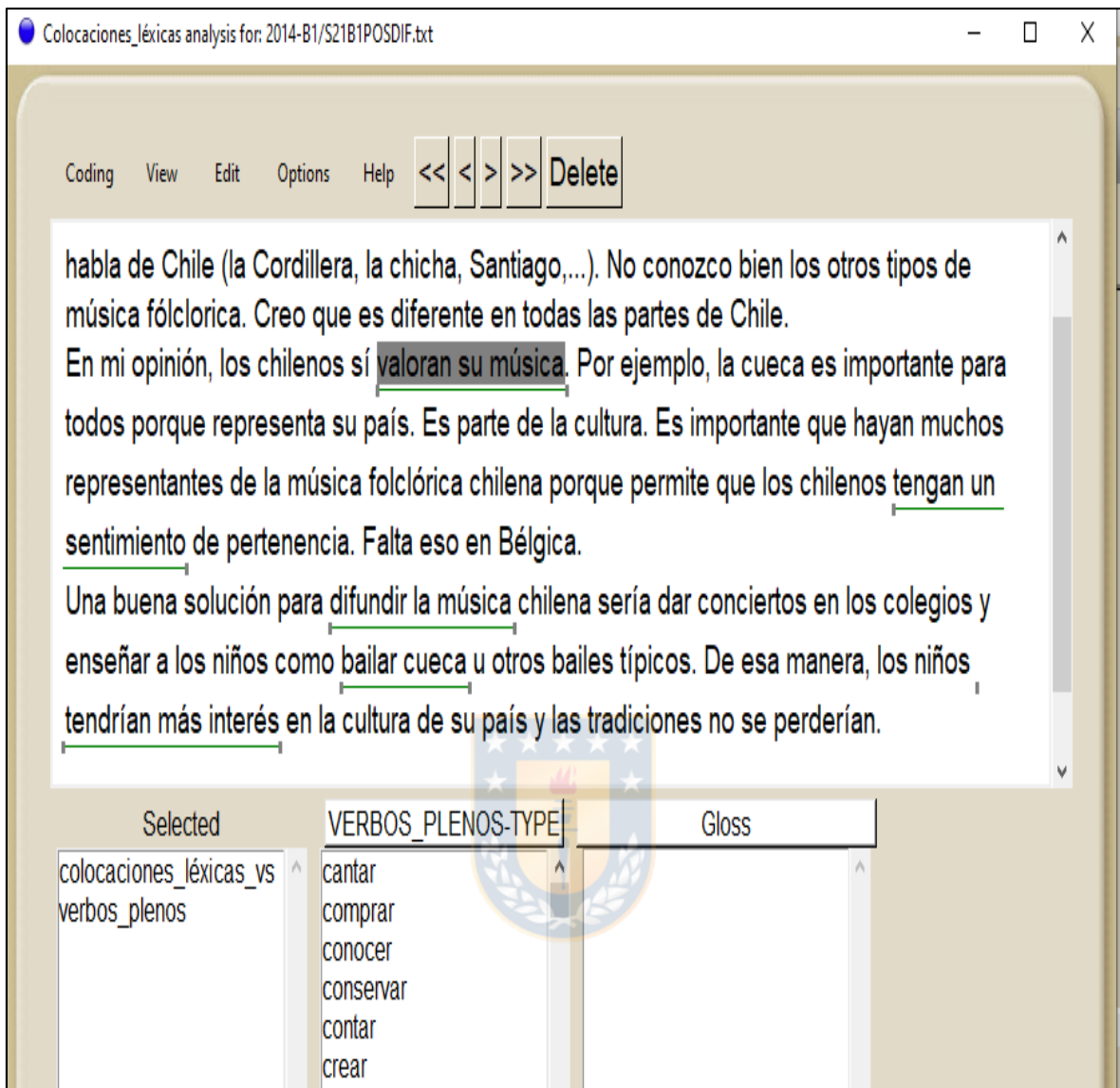


Figura 2. Aplicación de taxonomía.

6.7. Criterios para la identificación de colocaciones léxicas simples verbo-nominales

Hasta la fecha, los estudios centrados en la producción de colocaciones léxicas tanto en lengua materna (Martos, 2016) como en ELE (Pérez, 2014; García, 2017; García y Ramos, 2018) han utilizado como criterios de identificación las características de estas combinaciones (Koike, 2001), vale decir, coocurrencia

frecuente, restricciones combinatorias, composicionalidad formal, vinculación de lexemas, relación típica entre los componentes y precisión semántica de la combinación. Para efectos de esta investigación, se han propuesto algunos criterios de identificación de colocaciones léxicas simples verbo-nominales.

Para determinar si una producción era una colocación, se utilizó como primer criterio que el verbo tuviera naturaleza transitiva, esto era, que la acción se proyectara en un sustantivo o grupo nominal. Ese componente tradicionalmente desde el enfoque gramatical se conoce como objeto o complemento directo, en el cual el verbo exige la presencia de un nombre. El diccionario de la RAE (2014) lo define como: "Nombre, pronombre, sintagma o proposición en función nominal, que completa el significado de un verbo transitivo". Por ejemplo, "hacer un viaje", "leer un libro", "escuchar música", "tener apoyo", "dar importancia", etc. Para ello, se detectaron todos los casos de complemento directo.

Un segundo criterio se basó en el procedimiento estadístico de frecuencia, propuesto por Firth, (1957), quien postuló que cualquier par de palabras que co-aparezca al menos dos veces en un corpus es una colocación potencial; sin embargo, es mejor observar una mayor frecuencia; por ejemplo, un mínimo de tres, cinco o incluso diez de coocurrencia. La frecuencia de coaparición "se trata simplemente del cómputo de veces que una determinada combinación se da en un corpus concreto" (García, 2017, p.26).

Para García (2017), la frecuencia desempeña una función relevante en la selección de las unidades léxicas y también la intuición por parte del docente. Por lo tanto, se seleccionaron como colocaciones verbo-nominales con verbos de apoyo todas las combinaciones que se repitieran al menos tres veces en el sub-corpus lingüístico al que pertenecían. La elección del criterio de frecuencia se justifica, dado que permite la asignación de colocaciones a distintos niveles de competencia comunicativa (Ferrando, 2012, p. 360-361) y la definición de fraseologismos (Nation, 2001, p. 329; Martínez, 2013). Al respecto, la frecuencia de los usos que presentan determinadas combinaciones léxicas en los corpus lingüísticos permiten analizar las producciones escritas de aprendientes de

segundas lenguas y compararlas con producciones de nativos o con las de aprendices de distintos niveles (Durrant y Schmitt, 2009; Bestgen y Granger, 2014; Granger y Bestgen, 2014; Lorenz, 1999)

Un tercer criterio fue considerar no solo las formas conjugadas del verbo, sino también las formas no personales, es decir, casos de colocaciones, en las que el colocado era un “infinitivo” o un “gerundio”. Por ejemplo, “mantener tradiciones” o “manteniendo tradiciones”. En estos dos casos, se contabilizaron ambas estructuras como colocaciones de V+S.

6.8. Revisión textual por parte de otros etiquetadores

Con la finalidad de validar la identificación de este grupo de colocaciones, estas fueron sometidas a un segundo análisis por parte de cinco hablantes nativos, quienes etiquetaron de forma manual el 100% de ambos sub-corpus lingüísticos. Esta etapa se justifica metodológicamente en los procedimientos postulados tanto por Granger (2004) como por Jarvis y Pablenko (2007), quienes establecen la comparación entre distintos corpus lingüísticos de aprendientes en formato electrónico.

En este contexto, se compara la problemática lingüística de las colocaciones léxicas simples verbo-nominales en la L2 de los aprendientes, es decir, por niveles de competencia del ELE y en la L1 de los hablantes nativos de la lengua objeto de estudio, esto es, español como LM.

Estos etiquetadores eran estudiantes de pregrado de la carrera Pedagogía en Educación Media en Lenguaje y Comunicación de una universidad chilena y que se encontraban cursando el último año académico. A cada uno de ellos se les solicitó completar una ficha con sus datos y se les explicó en una sesión de entrenamiento el procedimiento que debían realizar mediante ejemplos (*ver Anexo 13.1*). Cada colaborador debió: 1) detectar todos los casos en que el estudiante de ELE produjo casos de complementos directos, ya sea correctamente o

incorrectamente (agramaticales o inexistentes en español), 2) subrayar el segmento que contenía dicho complemento y 3) señalar si consideraba su uso frecuente o poco frecuente en el español (F/PF).



CAPÍTULO III. RESULTADOS Y DISCUSIÓN

Este apartado tiene por objetivo describir exhaustivamente los resultados obtenidos del análisis de colocaciones léxicas simples con verbos de apoyo y verbos de significado pleno, a partir de los datos procesados en el sub-corpus CAELE. Asimismo, tiene por objetivo discutir los resultados a la luz de la literatura científica. Su presentación se divide en estudio 1 y estudio 2, y se organiza principalmente en gráficos y tablas.

7. Resultados Estudio 1: Frecuencia colocacional en el sub-corpus CAELE

A continuación, se describen las frecuencias obtenidas considerando el análisis de 341 tareas de escritura. 188 textos correspondieron al nivel A2, mientras que 153, al nivel B1. El proceso de etiquetado fue realizado de forma manual con la taxonomía implementada en el *software*.

Estos resultados respondieron a los objetivos específicos, pertenecientes al primer objetivo general de investigación. Estos fueron: 1) Identificar las colocaciones léxicas simples verbo-nominales de competencia del ELE A2 y B1 utilizando los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico, 2) Clasificar las colocaciones léxicas simples verbo-nominales en construcciones con verbos de apoyo y en construcciones con significado pleno y 3) Delimitar las frecuencias de bases y colocados de las estructuras en estudio. También, los hallazgos respondieron a la H1: La producción de colocaciones léxicas simples con verbos de apoyo es menor que la producción con verbos de significado pleno.

7.1. Colocaciones léxicas verbo-nominales

El análisis colocacional realizado permitió evidenciar un total de 513 colocaciones. Tal como se muestra en el Gráfico 1, la mayoría correspondió a verbos con significado pleno (464), mientras que un número menor estuvo constituido por verbos de apoyo (49). En el primer grupo destacaron colocados como “difundir” (“difundir una leyenda”), “escuchar” (“escuchar música”) y “valorar” (valorar la cultura). Al dividir las 513 colocaciones encontradas por el total de 341 tareas de escritura, se puede afirmar que cada aprendiente produce en promedio 1,5 colocaciones léxicas verbo-nominales por producción textual.

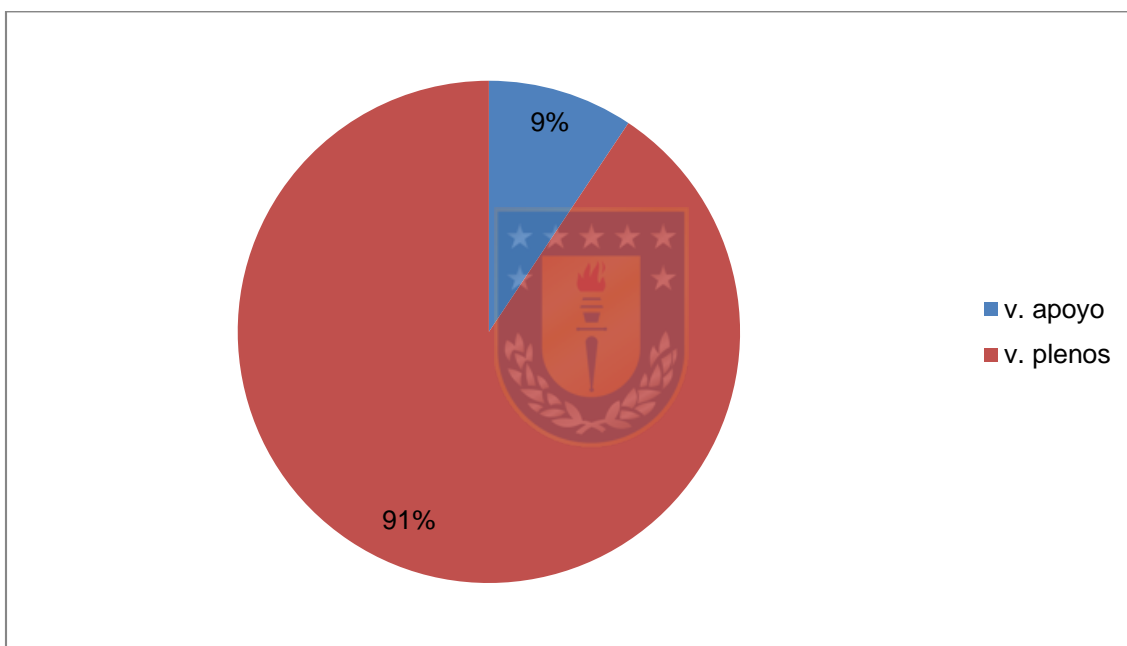


Gráfico 1. Colocaciones léxicas V+S en el sub-corpus CAELE. Elaboración propia.

7.2. Colocaciones léxicas más frecuentes con verbos de apoyo

En esta tabla se aprecian las colocaciones con verbos de apoyo presentes en el sub-corpus CAELE. En este tipo de colocados se registraron 49 casos. Destacó la colocación “tener importancia”, producida por los aprendientes. “Tener” fue el

colocado más frecuente con 37 usos, mientras que importancia, la base más frecuente con 24 usos (ver Tabla 14).

N°	Colocaciones VA	Fr.	%
1	tener importancia	15	30,61%
2	dar importancia	9	18,37%
3	tener influencia	7	14,29%
4	tener base	4	8,16%
5	tener interés	4	8,16%
6	tener acceso	4	8,16%
7	hacer una pintura	3	6,12%
8	tener origen	3	6,12%
Total	-	49	100%

Tabla 14. Colocaciones léxicas V+S con verbos de apoyo en el sub-corpus CAELE. Elaboración propia.

Se puede señalar incipientemente que no existió una correspondencia directa o evidente entre la mayoría de las bases (“importancia”, “influencia”, “base”, “interés”, “acceso” y “origen”) y los temas textuales sobre los cuales escribieron los aprendientes de ELE. La base “pintura” entregó noción de que los aprendientes redactaron una composición sobre alguna temática artística, dicho tópico fue “Pinturas costumbristas”.

7.3. Colocaciones léxicas más frecuentes con verbos plenos

En la Tabla 15 se presentan las 15 colocaciones léxicas más frecuentes en el sub-corpus de aprendientes. Este grupo de colocaciones representó el 57% de todas las combinaciones con verbo de significado pleno identificadas en las tareas de escrituras (ver Anexo 11.2). Es relevante observar que algunas bases se encontraban modificadas por artículos determinados e indeterminados, por ejemplo, “tener un libro” o “valorar la artesanía”, mientras que otras carecían de modificadores. Asimismo, se puede observar que las bases se presentaron como sustantivos en algunos casos con número singular y en otros casos, con número plural.

N°	Colocaciones VP	Fr.	%
1	escuchar música	73	27,55%
2	valorar la música	36	13,58%
3	valorar la pintura	28	10,57%
4	valorar la artesanía	22	8,30%
5	valorar la literatura	14	5,28%
6	comprar cosas	12	4,53%
7	mantener una tradición	12	4,53%
8	leer un libro	11	4,15%
9	preferir la música	9	3,40%
10	escuchar canciones	9	3,40%
11	escuchar reguetón	8	3,02%
12	hacer cosas	8	3,02%
13	tener historia	8	3,02%
14	vender productos	8	3,02%
15	mantener la cultura	7	2,64%
Total	-	265	100,00%

Tabla 15. Colocaciones léxicas V+S con verbos de significado pleno en el sub-corpus CAELE. Elaboración propia.



7.3.1. Bases más frecuentes de las colocaciones léxicas verbo-nominales con verbos de significado pleno

Las bases que conformaron las 464 colocaciones con verbos plenos correspondieron a 29 (*ver Tabla 16*). La que predominó la producción colocacional fue el sustantivo “música” (29, 96%), esto quiere decir, que fue capaz de elegir distintos colocados, por ejemplo, “escuchar”, “valorar” y “preferir” y combinarse en una estructura. La mayoría de las bases fueron sustantivos abstractos, por ejemplo: “cultura” (6, 03%), “canciones” (4, 53%), “historias” (4,31%), “tradición” (3,23%), “atención” (1, 08%), “idioma” (0, 65%) y “tiempo” (0, 65%).

N°	Bases	Frecuencia	%
1	música	139	29,96%
2	pintura	34	7,33%
3	artesanía	32	6,90%
4	cultura	28	6,03%
5	cosas	27	5,82%
6	canciones	21	4,53%
7	historias	20	4,31%
8	literatura	18	3,88%
9	productos	17	3,66%
10	libro	17	3,66%
11	tradición	15	3,23%
12	instrumentos	12	2,59%
13	leyendas	9	1,94%
14	parte	9	1,94%
15	reguetón	8	1,72%
16	ganas	7	1,51%
17	pena	6	1,29%
18	poema	5	1,08%
19	dinero	5	1,08%
20	atención	5	1,08%
21	cuenta	4	0,86%
22	premio	4	0,86%
23	temas	4	0,86%
24	texto	3	0,65%
25	opinión	3	0,65%
26	idioma	3	0,65%
27	lugar	3	0,65%
28	tiempo	3	0,65%
29	pueblo	3	0,65%
Total	-	464	100,00%

Tabla 16. Bases en colocaciones léxicas simples V+S. Elaboración propia.

A diferencia de las bases con verbos de apoyo, estas 29 bases establecieron una relación de correspondencia léxica directa con los centros temáticos sobre los cuales los aprendientes de ELE produjeron sus tareas de escritura. Por ejemplo, la base “música” se correspondió con los temas: a) Música latinoamericana, b) Música popular chilena, c) Música juvenil, d) Música folclórica del país de origen y e) Música folclórica chilena. De igual forma, la base “canciones”, “instrumentos” y “reguetón” se circunscribió en los temas anteriores.

7.3.2. Colocados más frecuentes de las colocaciones léxicas verbo-nominales con verbos de significado pleno

En el Gráfico 2 se da cuenta de los 10 colocados más frecuentes (348), utilizados por los aprendientes del sub-corpus. Destacó el verbo “valorar” (31%), “escuchar” (26%) y “leer” (7%). De este último colocado, se puede afirmar que fue más frecuente la combinación léxica “leer un libro” que “escribir un libro”.

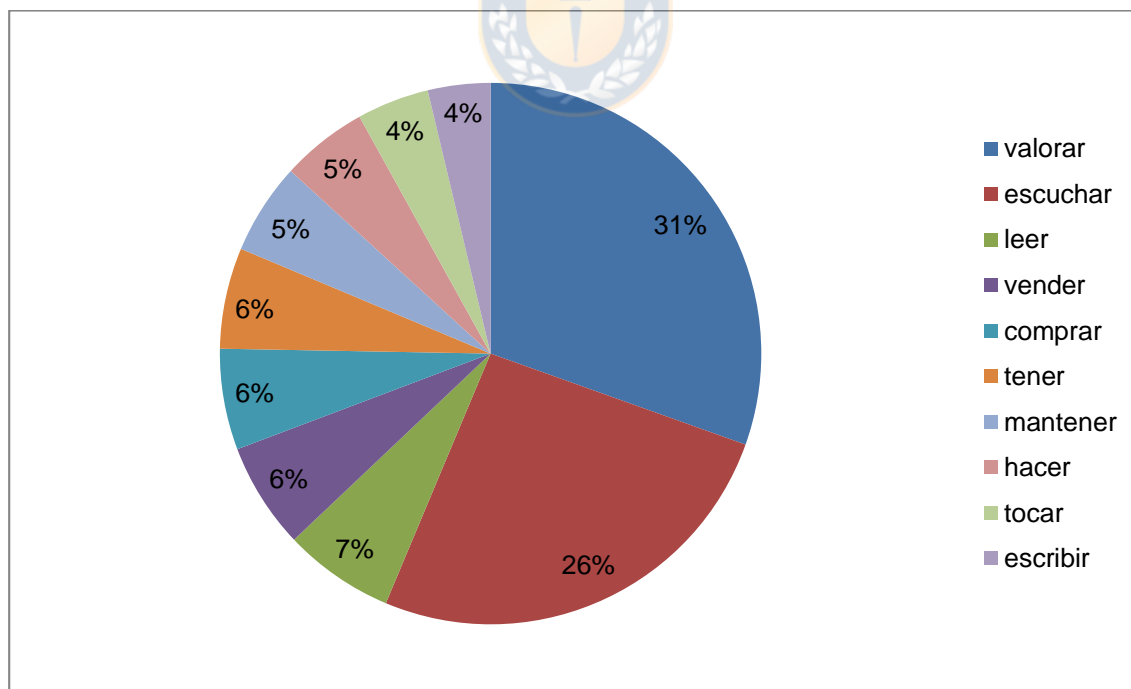


Gráfico 2. Colocados en colocaciones léxicas simples V+S. Elaboración propia.

7.4. Prueba estadística *t-test*

Con la finalidad de establecer si la diferencia producida entre los dos tipos de colocaciones estudiadas era significativa, se ejecutó la prueba *t-test* con un intervalo de confianza del 95%. En la Tabla 17, los resultados indican que el promedio fue mayor en la categoría de colocaciones con verbos de significado pleno en relación a las colocaciones con verbos de apoyo. En cuanto a la desviación estándar, esta medida se presentó muy alta en las colocaciones con verbos plenos, esto quiere decir que, por una parte, hubo aprendientes que produjeron con alta frecuencia determinadas colocaciones de esta naturaleza y, por otra parte, aprendientes que produjeron este tipo de colocaciones con una frecuencia mínima. La idea anterior lleva a afirmar una alta dispersión de colocaciones con estos verbos. La misma medida se presentó baja en las colocaciones con verbos de apoyo. Finalmente, la prueba arrojó un valor $P=0.001$, el cual se considera estadísticamente significativo.

Grupo	Colocaciones con verbos de apoyo	Colocaciones con verbos de significado pleno
Valor de P	P= 0,0001	
Promedio	0,80	7,61
Desviación estándar	2,52	10,32
N° de colocaciones	49	464

Tabla 17. *T-test* para grupos pareados en relación a tipos de colocaciones léxicas verbo-nominales. Elaboración propia.

Los resultados obtenidos tienen gran relevancia, dado que logran responder positivamente a la H1: La producción de colocaciones léxicas simples con verbos de apoyo es menor que la producción con verbos de significado pleno, la cual ha sido formulada en esta investigación.

7.5. Colocaciones léxicas por nivel de competencia comunicativa

7.5.1. Colocaciones léxicas verbo-nominales más frecuentes en el nivel de competencia A2

En el Gráfico 3 se observa la frecuencia de las colocaciones con verbos de apoyo y la frecuencia de estas combinaciones con verbos plenos producidas por los aprendientes de nivel inicial. Se puede notar que la primera categoría de colocaciones fue poco frecuente (22) en relación a las estructuras en el cual el significado fue pleno (248).

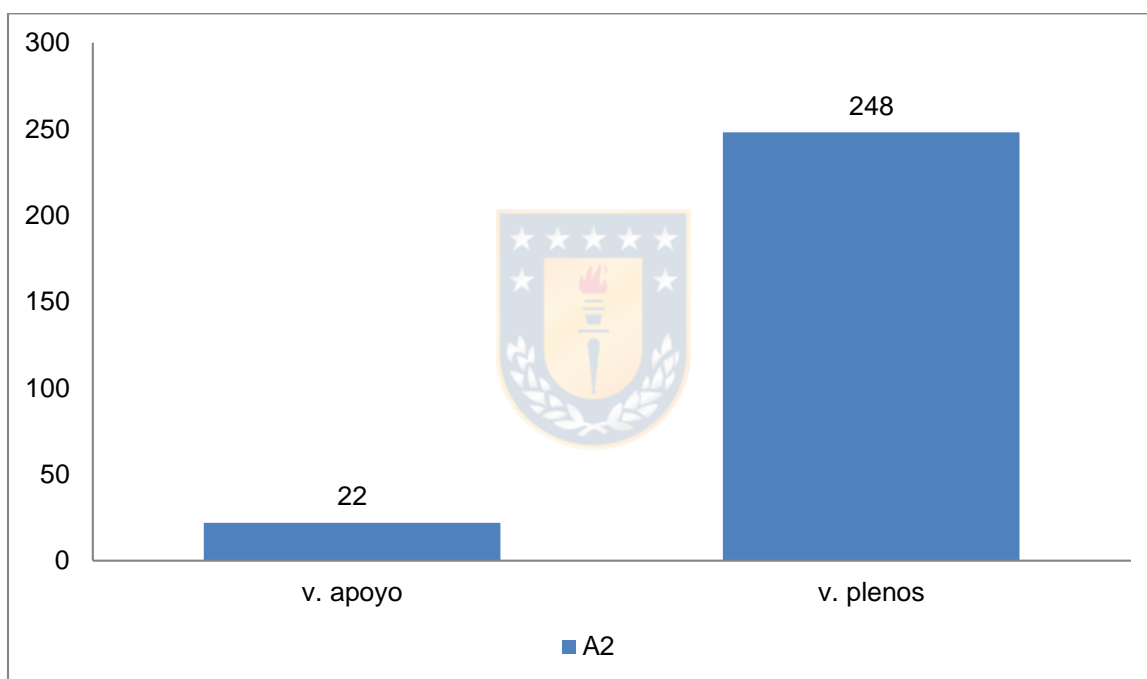


Gráfico 3. Colocaciones léxicas verbo-nominales en nivel A2. Elaboración propia.

7.5.2. Colocaciones léxicas verbo-nominales más frecuentes en el nivel de competencia B1

En el nivel de competencia B1, los resultados de las colocaciones léxicas verbo-nominales fueron similares. Las combinaciones con verbos de apoyo fueron

menores (27) en comparación a las combinaciones con verbos de significado pleno (216) (ver Gráfico 4).

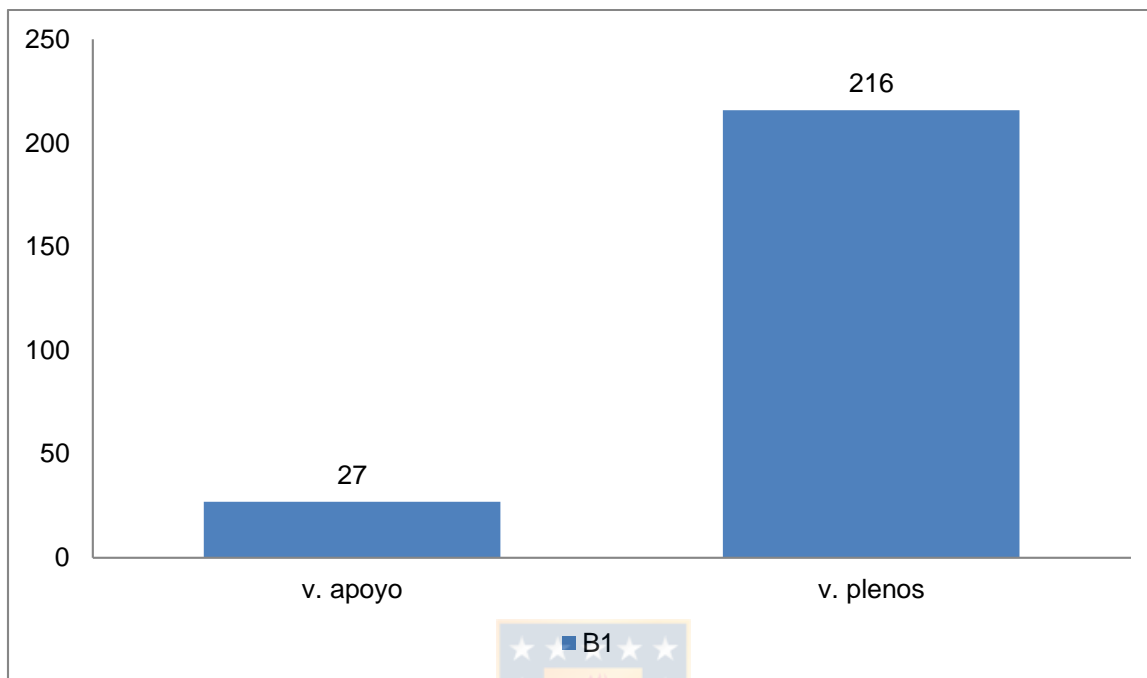


Gráfico 4. Colocaciones léxicas verbo-nominales en nivel B1. Elaboración propia.

En las siguientes tablas, se detalla los tipos de colocaciones léxicas simples verbo-nominales identificadas por nivel y se muestra algunos ejemplos que permiten observar el funcionamiento de esta problemática lingüística.

7.6. Colocaciones léxicas verbo-nominales con verbo de apoyo más frecuentes por nivel de competencia comunicativa

7.6.1. Colocaciones léxicas verbo-nominales con verbos de apoyo más frecuentes en el nivel de competencia A2

En la Tabla 18 se da cuenta sobre la producción de colocaciones léxicas con verbos de apoyo en el nivel de competencia comunicativa A2. Los aprendientes produjeron 7 tipos de combinaciones. Se destacó la colocación con verbo soporte

“tener influencia” (31, 82%), que puede ser sustituida por la acción verbal “influenciar”.

N°	A2	Fr.	%
1	tener influencia	7	31,82%
2	tener base	3	13,64%
3	tener origen	3	13,64%
4	hacer pintura	3	13,64%
5	tener importancia	3	13,64%
6	tener acceso	2	9,09%
7	tener interés	1	4,55%
Total		22	100%

Tabla 18. Colocaciones con verbos de apoyo en nivel A2. Elaboración propia.

Es relevante, ilustrar algunos casos colocacionales²⁹ con la finalidad de mostrar el contexto de uso en el que se produjeron estas combinaciones. Conviene observar los siguientes ejemplos:

- 1) “*Ántes la musica alemán tenía grande influencia, en el mundo de la música clasica*” (aprendiente 2015S20A2),
- 2) “*En el suroeste, tenemos una influencia fuerte de la música mexicana, de latinpop, y de los campesinos*” (aprendiente 2015S13A2),
- 3) “*La literatura de EEUU tiene su base en políticas y correspondencias al medio de los colonistas y Britancia*” (aprendiente 2014S11A2) y
- 4) “*La baila nacional de Chile es la cueca, que tiene sus orígenes en la música española, africana, peruviiana y la musica indígena de Chile*” (aprendiente 2014S8A2).

²⁹ Todos los casos colocacionales ilustrados han sido extraídos tal cual fueron producidos en las tareas de escritura. Por esta razón, algunos contienen errores léxicos y gramaticales. No han sido corregidos con el objetivo de mostrar su contexto auténtico de producción.

7.6.2. Colocaciones léxicas verbo-nominales más frecuentes con verbos de apoyo en el nivel de competencia B1

Los aprendientes de este nivel intermedio produjeron 5 tipos de colocaciones léxicas v+s (ver Tabla 19), pero con menor variedad léxica a diferencia de los aprendientes de nivel A2. Se destacó el par de colocaciones “tener importancia” (48, 15%) y “dar importancia” (29, 63%). Ambas con la misma base, pero con distinto colocado. Se puede agregar que el significado que le otorga el aprendiente de ELE es idéntico si se considera el colocado como des-lexicalizado o parcialmente carente de significado en cuanto a su información semántica, por lo que el significado estaría presente en la base de estas combinaciones léxicas simples.

N°	B1	Fr.	%
1	tener importancia	13	48,15%
2	dar importancia	8	29,63%
3	tener interés	3	11,11%
4	tener acceso	2	7,41%
5	tener base	1	3,70%
Total		27	100%

Tabla 19. Colocaciones con verbos de apoyo en nivel B1. Elaboración propia.

En este nivel intermedio fueron seleccionados los siguientes casos colocacionales:

- 1) “Anteriormente los jóvenes no dieron importancia a estos músicos” (aprendiente 2014S20B1),
- 2) “En las ferias u otros lugares donde personas venden su artesanía, muchos chilenos tienen interés en los productos” (aprendiente 2015S24B1),
- 3) “La literatura tenía gran importancia para la unión de los chicos países alemanes y la formación de una nación” (aprendiente 2015S11B1) y

- 4) “La gente inglesa le da mucha importancia a los poetas y narradores ingleses” (aprendiente 2014S19B1).

7.7. Colocaciones léxicas con verbos de significado pleno por nivel de competencia comunicativa

7.7.1. Colocaciones léxicas verbo-nominales más frecuentes con verbos de significado pleno en el nivel de competencia A2

En la Tabla 20, se observan las 10 colocaciones más frecuentes con verbos de significado pleno de un total de 248 (ver Anexo 11.2). Se puede señalar que en el nivel A2 se concentró una cantidad de 125. La colocación más frecuente fue “escuchar música”. Su base apareció en tres tipos de colocaciones “escuchar música”, “valorar música” y “preferir música”.

N°	A2	Fr.	%
1	escuchar música	35	28,00%
2	valorar la música	21	16,80%
3	valorar la pintura	16	12,80%
4	valorar la artesanía	9	7,20%
5	comprar cosas	8	6,40%
6	mantener la tradición	8	6,40%
7	tener historia	8	6,40%
8	preferir la música	7	5,60%
9	valora la literatura	7	5,60%
10	usar instrumentos	6	5%
Total	-	125	100%

Tabla 20. Colocaciones con verbos plenos en nivel A2. Elaboración propia.

Los siguientes casos colocacionales fueron extraídos de los textos producidos por los aprendientes del nivel A2.

- 1) “En mi casa siempre escucho música, por que esa haceré feliz. Especialmente música electronica y música de rock, pero a veces escucho música distinta tambien” (aprendiente 2015S12A2),
- 2) “Los chilenos valoran sus pinturas porque es la forma de expresión de las artistas” (aprendiente 2016S32A2),
- 3) “Para mentener la tradiciones pienso que debemos poner más énfasis en la cultura de los pueblos originarios” (aprendiente 2015S17A2) y
- 4) “Entonces yo quiero comprar cosas bonitas para mi familia” (2016S30A2).

7.7.2. Colocaciones léxicas verbo-nominales más frecuentes con verbos de significado pleno en el nivel de competencia B1

En la Tabla 21, se observan los 10 tipos de tipos de colocaciones más frecuentes en el nivel B1 de un total de 216 (ver anexo 11.3). “Escuchar música” (32,48%) fue la combinación más utilizada por este grupo de aprendientes, a diferencia de “escuchar canciones” (5,13%), la cual fue usada solo 6 veces en sus producciones textuales. Al parecer, este grupo de ELE presenta una preferencia por la selección semántica de la base. Prefieren utilizar “música” antes que “canciones”, aunque ambas construcciones expresen un sentido similar.

N°	B1	Fr.	%
1	escuchar música	38	32,48%
2	valorar la música	15	12,82%
3	valorar la artesanía	13	11,11%
4	valorar la pintura	12	10,26%
5	leer un libro	7	5,98%
6	valorar la literatura	7	5,98%
7	vender artesanía	7	5,98%
8	comprar productos	6	5,13%
9	escuchar canciones	6	5,13%
10	vender productos	6	5,13%
Total	-	117	100%

Tabla 21. Colocaciones con verbos plenos en nivel B1. Elaboración propia.

Algunos ejemplos extraídos fueron:

1) *“También Tuve que leer un libro que escribió mi profesora de francés del colegio pero ella no es famosa”* (2014S2B1),

2) *“En las ferias de turistas deberían vender mas artesanía chilenos y menos productos hecho en China”* (2015S9B1),

3) *“En Europea todos los jóvenes escuchan muchas canciones “comerciales”* (2014S13B1) y

4) *“Creo que los chilenos valoran sus artesanías”* (2016S41B1).

7.8. Bases más frecuentes de colocaciones léxicas con verbos de apoyo por nivel de competencia comunicativa

7.8.1. Bases más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia A2

En el Gráfico 5 se puede observar las bases más frecuentes que utilizaron los aprendientes de nivel A2. Solo fueron 7 combinaciones, esto se debió a que las colocaciones de este tipo fueron limitadas en relación a las conformadas con verbos plenos. La más frecuente fue “influencia” (32%), mientras que la menos productiva, “interés” (4%). La mayoría de las bases pudieron ser definidas como sustantivos abstractos, a excepción de la base “pintura” (14%) que es un nombre concreto.

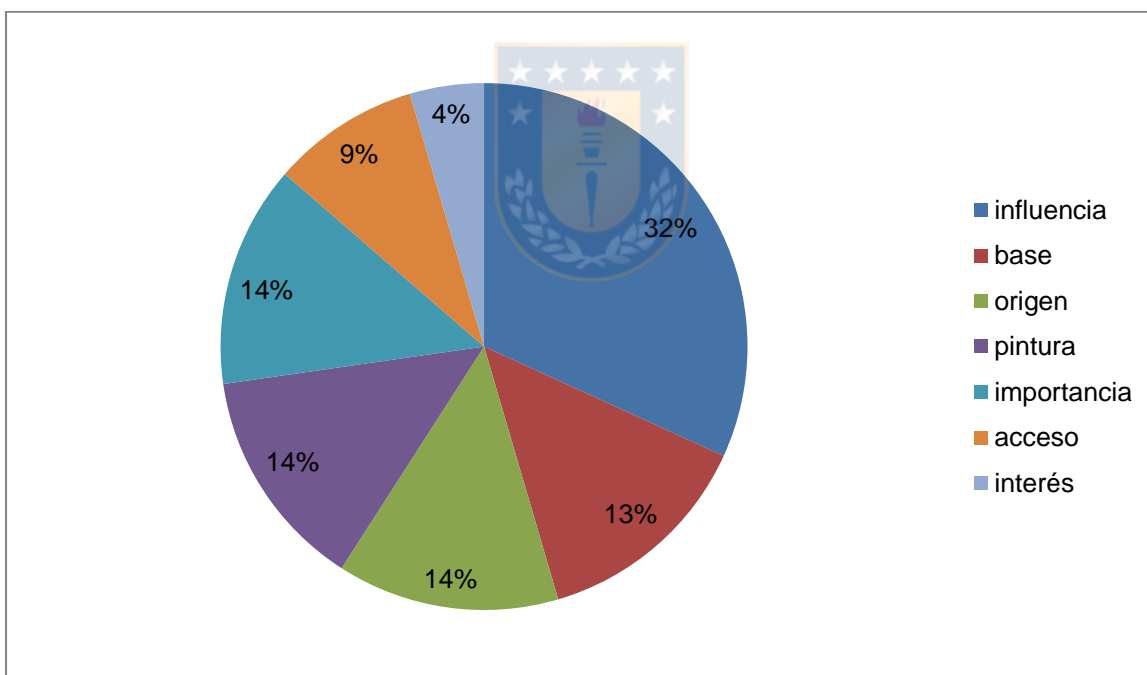


Gráfico 5. Bases con verbos de apoyo en nivel A2. Elaboración propia.

7.8.2. Bases más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia B1

En el Gráfico 6 se muestran las bases seleccionadas por los aprendientes de nivel B1. El término “importancia” (78%) fue el más usado por este grupo. Estos sustantivos cumplieron el rol de establecer un vínculo léxico-semántico con el verbo y así formar la combinación. Llama la atención que todos estos nombres fueron clasificados como abstractos. Al parecer las bases con esta naturaleza manifiestan la tendencia de seleccionar verbos soporte, ya que la misma relación se estableció en el nivel A2. Asimismo, estas bases no presentan evidencia o alguna huella léxica que los vincule a alguna temática específica.

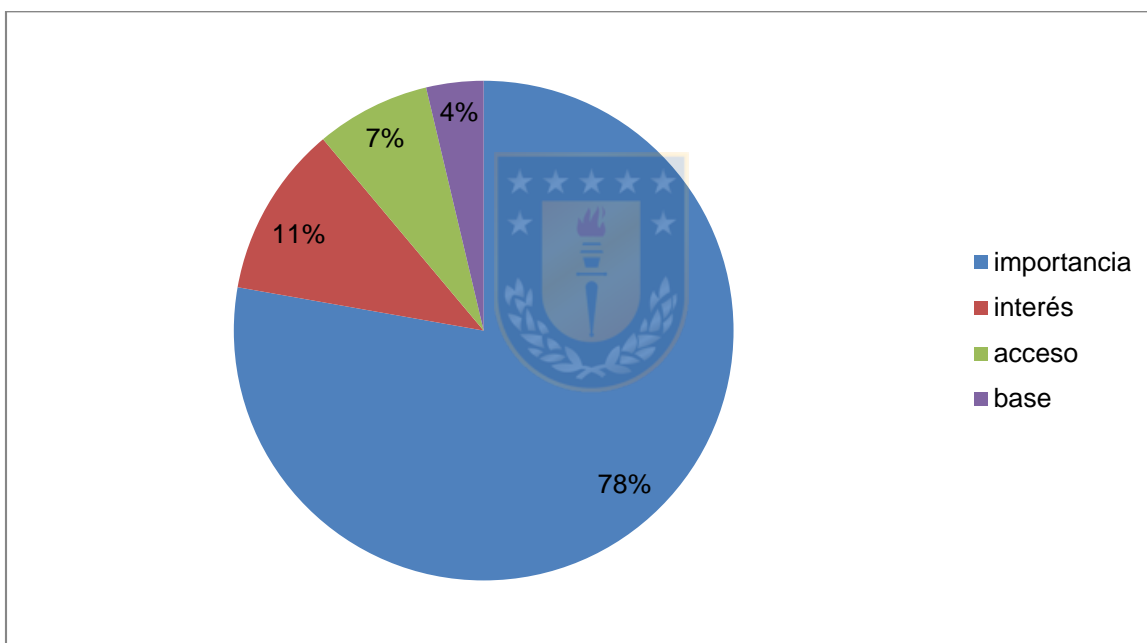


Gráfico 6. Bases con verbos de apoyo en nivel B1. Elaboración propia.

7.9. Colocados más frecuentes de colocaciones léxicas con verbos de apoyo por nivel de competencia comunicativa

7.9.1. Colocados más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia A2

En el nivel A2, la selección de colocados se caracterizó por la predominancia del verbo “tener” (86%), es decir, la mayoría de las bases atrajeron a este tipo de colocado. Se constató un total de 19 verbos de este tipo. El colocado “hacer” (14%) fue el menos productivo por los aprendientes del nivel principiante de ELE. (ver Gráfico 7).

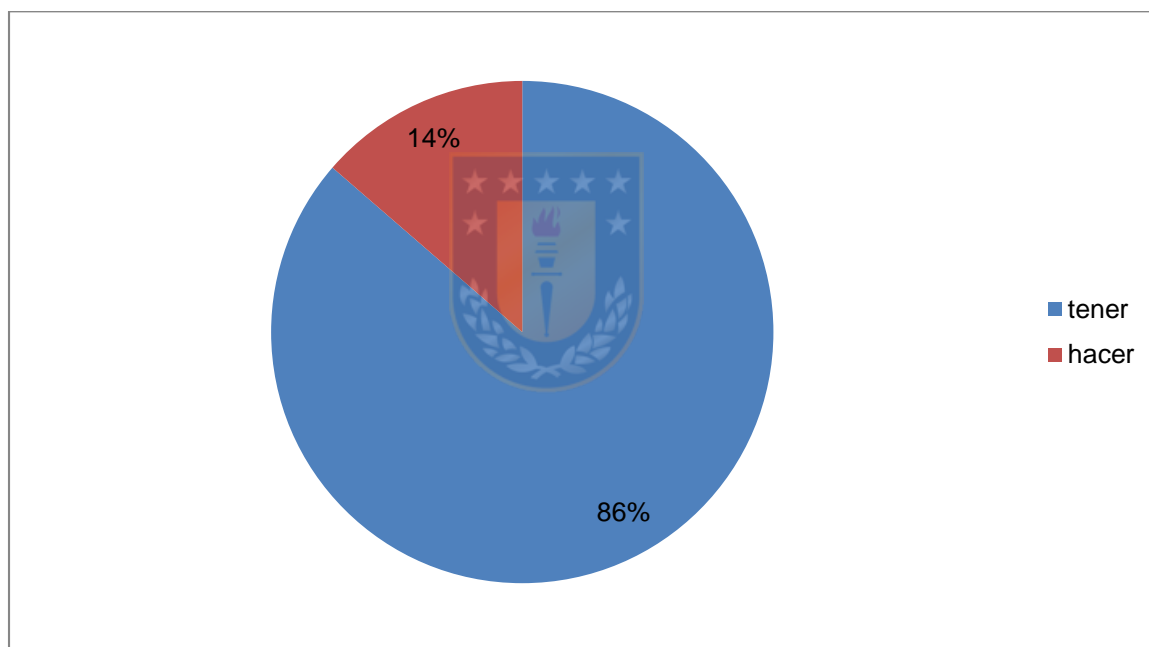


Gráfico 7. Colocados con verbos de apoyo en nivel A2. Elaboración propia.

Por ejemplo:

- 1) *“También hay gente quien hace pinturas o muñecas. Esas cosas pueden representar temas diferentes como la religión o los políticos, o quizás el medio ambiente”* (aprendiente 2015S3A2) y

- 2) *“Si yo estará una artesana, me gustaría participar en las ferias y hacer pinturas”* (aprendiente 2015S6A2).

7.9.2. Colocados más frecuentes de colocaciones léxicas con verbos de apoyo en el nivel de competencia B1

En cuanto a los colocados, en el Gráfico 8 se muestra que el verbo más usado correspondió a “tener” (70%), seguido del verbo “dar” (30%). Si bien se encontraron varios casos de posibles colocaciones con verbos de apoyo, estas no reunieron todos los criterios especificados en la sección de metodología, en consecuencia, hubo una disminución de estas estructuras.

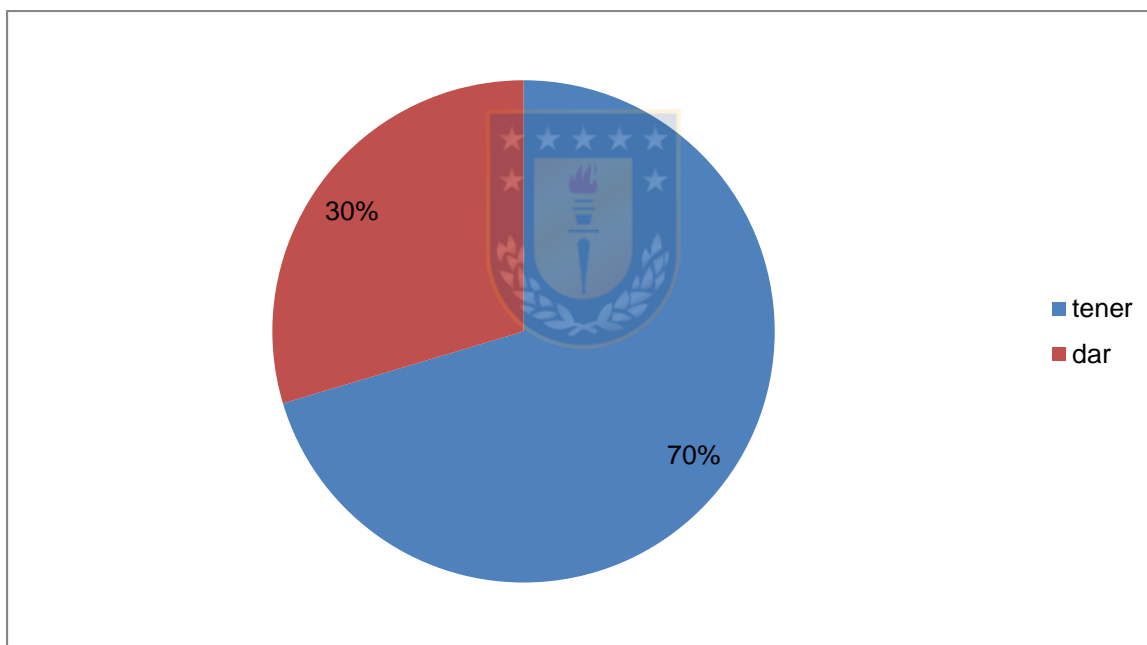


Gráfico 8. Colocados con verbos de apoyo en nivel B1. Elaboración propia.

Algunos ejemplos producidos por estos aprendientes fueron:

- 1) *“Tal vez se debería realizar algunas acciones para ayudar a los artesanos a difundir sus productos. Una posibilidad sería venderlos también venderlos en línea para que la gente de otros países tenga acceso a la artesanía maravillosa de Chile”* (aprendiente 2014S20B1) y

2) “En Alemania, la gente da mucha importancia a los narradores y poetas de sus país” (aprendiente 2015S2B1).

7.10. Bases más frecuentes de colocaciones léxicas con verbos de significado pleno por nivel de competencia comunicativa

7.10.1. Bases más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia A2

En el Gráfico 9 se observan los 10 nodos más frecuentes de las colocaciones léxicas con verbos plenos producidas por los aprendientes de nivel inicial. Estos nodos conforman un total de 195. La base más productiva fue “música”, seleccionada 67 veces, (37%) y la menos frecuente “instrumentos” (4%). Ambos sustantivos se vincularon con las temáticas alusivas a la música, al igual que el nodo “canciones” (6%).

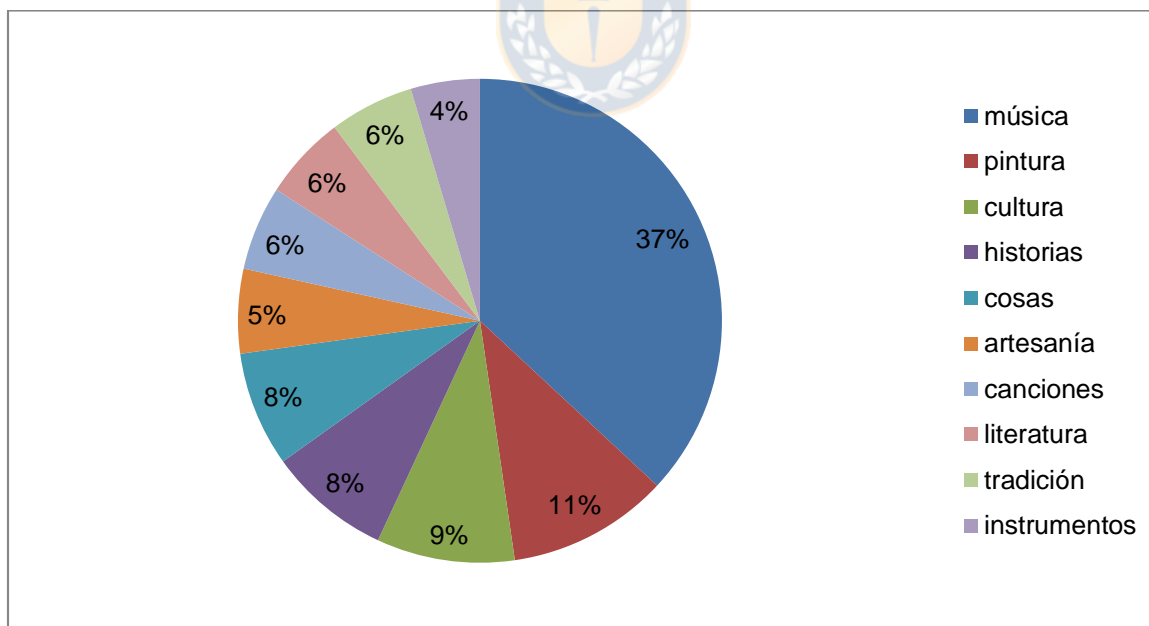


Gráfico 9. Bases con verbos plenos en nivel A2. Elaboración propia.

7.10.2. Bases más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia B1

En el Gráfico 10, se visualizan las 10 bases más frecuentes, las cuales alcanzan la suma de 170. Presentan dos características particulares: 1) la distribución es similar a las que se presentó en el nivel A2, 2) Siete de las bases seleccionadas (“música”, “artesanía”, “pintura”, “cosas”, “canciones”, “cultura” y “literatura”) por los aprendientes de este grupo coinciden con las del nivel inicial. “Música”, seleccionada 67 veces, (39%) se posicionó como la base más frecuente.

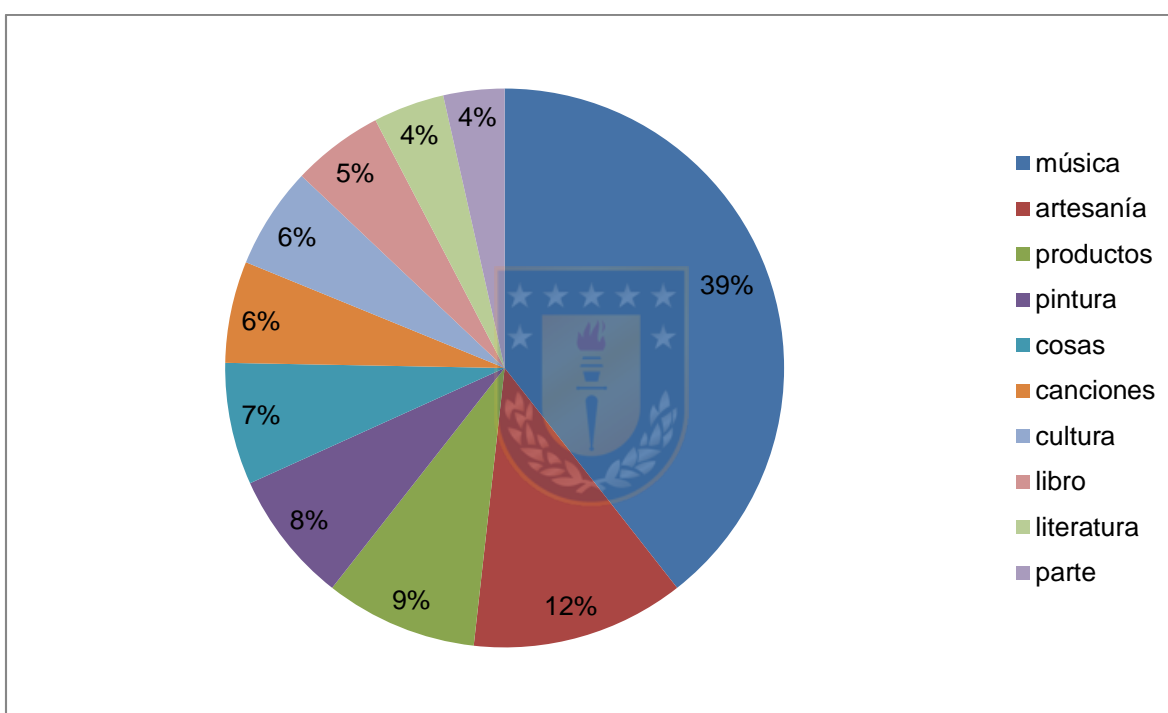


Gráfico 10. Bases con verbos plenos en nivel B1. Elaboración propia.

7.11. Colocados más frecuentes de colocaciones léxicas con verbos de significado pleno por nivel de competencia comunicativa

7.11.1. Colocados más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia A2

Los 10 tipos de colocados más frecuentes de nivel A2 fueron “valorar” (usado 47 veces) y “escuchar” (52 veces), los cuales representan en el Gráfico 11 más del 50% de los colocados. Ambos se vincularon léxicamente con la base “música” y formaron las colocaciones “escuchar música” y “valorar la música” que a su vez también presentaron alta frecuencia.

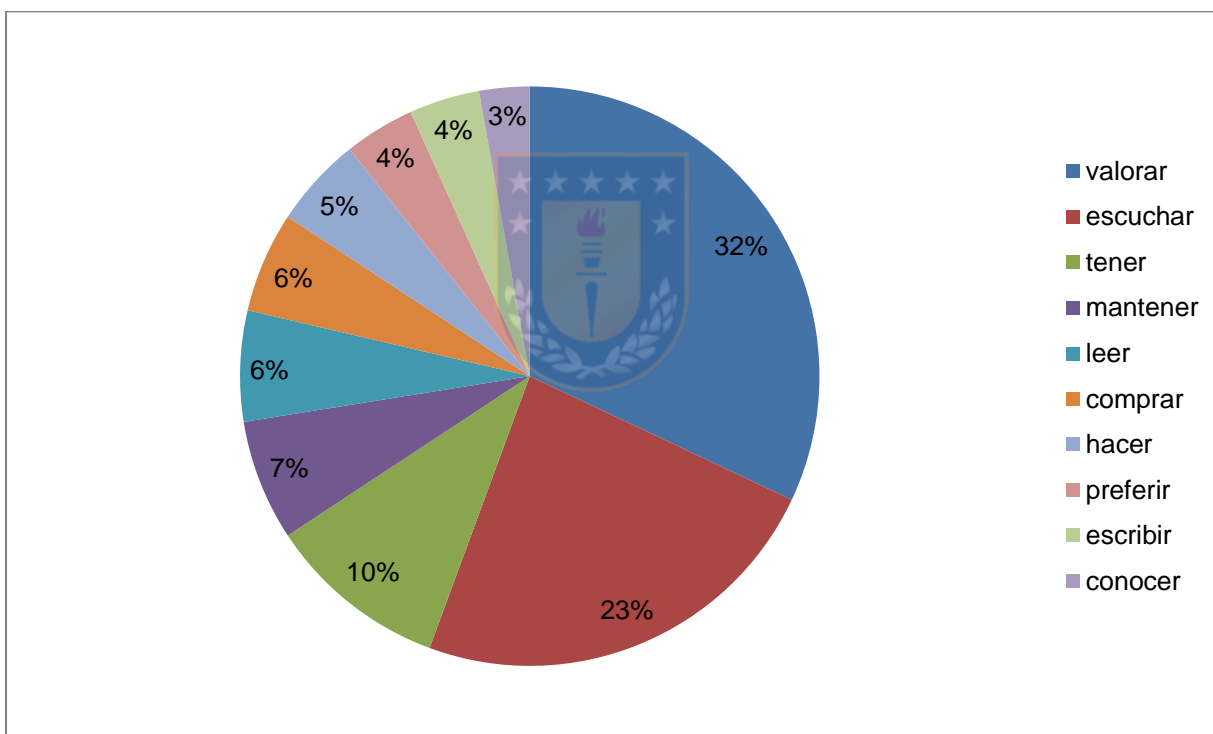


Gráfico 11. Colocados con verbos plenos en nivel A2. Elaboración propia.

7.11.2. Colocados más frecuentes de colocaciones léxicas con verbos de significado pleno en el nivel de competencia B1

En el nivel intermedio, se logra observar que la distribución de los 10 colocados más frecuentes en las colocaciones léxicas con verbos de contenido es similar a la que se presentó en el nivel A2. “Valorar” (49 usos) y “escuchar” (38 usos) fueron también los colocados que se posicionaron como más productivos.

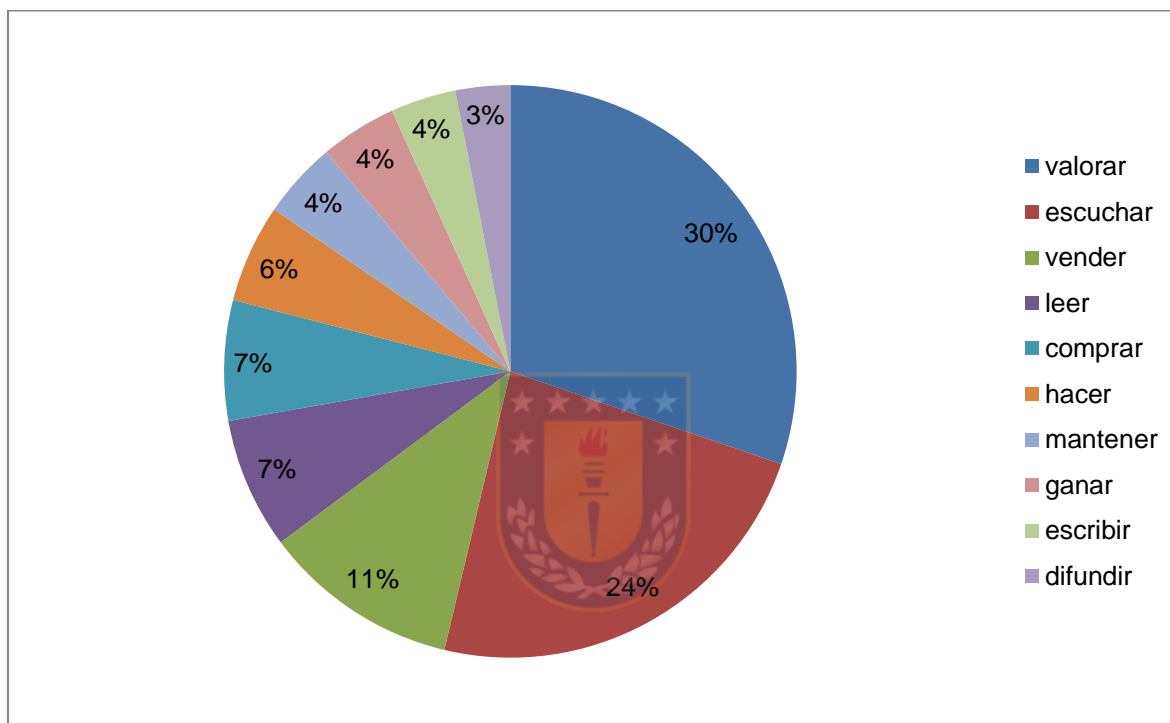


Gráfico 12. Colocados con verbos plenos en nivel B1. Elaboración propia.

Para concluir este estudio, los resultados obtenidos permiten asociar los hallazgos sobre los tipos de colocaciones léxicas verbo-nominales con la H1: La producción de colocaciones léxicas simples con verbos de apoyo es menor que la producción con verbos de significado pleno. En este sentido, se puede señalar a la luz de la aplicación del *t-test* que esta hipótesis ha sido confirmada, dado que se constató una diferencia significativa entre las colocaciones con verbos soporte y las conformadas con verbos plenos, en el sub-corpus CAELE. Asimismo, los resultados permiten describir exhaustivamente los usos colocacionales en los niveles de competencia estudiados. Los hallazgos responden de esta forma al

objetivo general 1) Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE.

8. Resultados Estudio 2: Comparación de la producción colocacional

Este estudio muestra los resultados de la comparación de colocaciones léxicas verbo-nominales. Los hallazgos respondieron a los siguientes objetivos específicos, correspondientes al segundo objetivo general Estos fueron: 1) Comparar las colocaciones léxicas simples verbo-nominales con verbos de apoyo de acuerdo con los niveles del ELE A2 y B1, 2) Comparar las colocaciones léxicas simples verbo-nominales con verbos plenos de acuerdo con los niveles del ELE A2 y B1, 3) Comparar las frecuencias de bases y colocados de las estructuras en estudio de acuerdo con los niveles del ELE A2 y B1. También los resultados dieron respuesta a la H2: La producción de colocaciones léxicas simples verbo-nominales es mayor en el nivel de competencia lingüística B1 que en el nivel A2.

Primero, se observa la comparación entre los dos niveles de competencia lingüística aplicando *t-tests* para muestras independientes. Segundo, se muestra la comparación entre las combinaciones objeto de estudio producidas por los aprendientes de niveles A2 y B1 con las combinaciones producidas por los hablantes nativos de español. Para ello, se utilizó la ejecución de las pruebas estadísticas ANOVA para evaluar la cantidad de colocaciones léxicas simples verbo-nominales.

8.1. Comparación de las colocaciones léxicas verbo-nominales en A2 y B1

En el siguiente gráfico se comparan las colocaciones verbo-nominales producidas por los 62 aprendientes del sub-corpus CAELE en los niveles A2 y B1. Por una parte, se puede observar que el nivel A2 produjo mayor cantidad de estas

combinaciones léxicas (ver Gráfico 13). Por otra parte, se puede señalar que ambos grupos de aprendientes produjeron un número similar de estas estructuras cuando utilizaron verbos de apoyo; sin embargo, con los verbos de significado pleno la distancia colocacional se hizo más notoria.

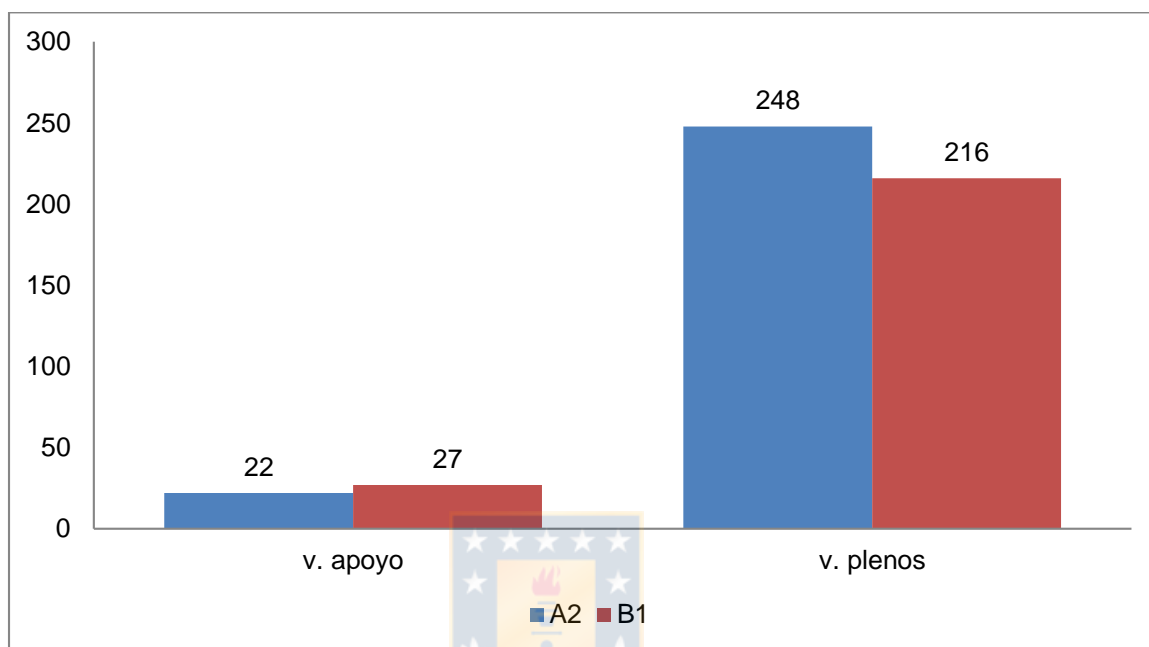


Gráfico 13. Colocaciones léxicas verbo-nominales en niveles de competencia lingüística. Elaboración propia.

8.2. Pruebas estadísticas *t-tests*

8.2.1. Colocaciones con verbos de apoyo en nivel A2 y B1

Para determinar si la diferencia de producción colocacional con verbos de apoyo fue estadísticamente significativa entre ambos niveles de competencia lingüística se ejecutó un *t-test* con un intervalo de confianza del 95%. En la Tabla 22, se muestran los hallazgos en relación a los 8 tipos de colocaciones léxicas verbo-nominales producidas por los aprendientes de ambos niveles de competencia lingüística (“tener influencia”, “tener base”, “tener origen”, “hacer pintura”, “tener” “importancia”, “tener acceso”, “tener interés” y “dar importancia”). El promedio fue mayor en el nivel intermedio (3,38), aunque la desviación estándar se elevó casi al

doble (4, 72), es decir, solo algunas categorías fueron productivas y otras no. La prueba arrojó un valor $P= 0.7335$, el cual no se considera estadísticamente significativo.

Grupo	A2	B1
Valor de P	P= 0,7335	
Promedio	2,75	3,38
Desviación estándar	1,91	4,72
Colocaciones	8	8
Nº de colocaciones	22	27

Tabla 22. *T-test* para grupos diferenciados en relación a colocaciones con verbos de apoyo. Elaboración propia.

8.2.2. Colocaciones con verbos plenos en nivel A2 y B1

A continuación, se visualizan si los resultados del uso de las colocaciones con verbos plenos fueron estadísticamente significativos entre el nivel inicial y el nivel intermedio. Esta prueba fue realizada con un intervalo de 95% de confianza en relación a los 61 tipos de colocaciones producidas con distintas frecuencias y comparadas en cada nivel de español (*ver Anexo 13.2 y 13.3*). Los resultados arrojaron una diferencia $P= 0,5911$, valor considerado como estadísticamente no significativo. El promedio de producción colocacional fue mayor en el nivel A2 (4,07) y la desviación estándar lo fue levemente en el nivel B1 (5,39), es decir, algunas colocaciones fueron altamente productivas por los aprendientes, mientras que otras tuvieron una escasa o nula producción, en consecuencia, se generó un mayor grado de dispersión (*ver Tabla 23*).

Grupo	A2	B1
Valor de P	P= 0,5911	
Promedio	4,07	3,54
Desviación estándar	5,36	5,39
Tipos de colocaciones	61	61
Nº de colocaciones por nivel	248	216

Tabla 23. *T-test* para grupos diferenciados en relación a colocaciones con verbos plenos. Elaboración propia.

Se pueden asociar los hallazgos obtenidos con la H2: La producción de colocaciones léxicas simples verbo-nominales es mayor en el nivel de competencia lingüística B1 que en el nivel A2. En este contexto, los resultados descritos permiten rechazar esta premisa, dado que se evidenció a través de dos pruebas *t-test* (la primera, para las colocaciones léxicas simples verbo-nominales con verbos de apoyo y la segunda, para las colocaciones léxicas simples verbo-nominales con verbos de significado pleno) que la producción de las colocaciones objeto de estudio fue mayor en el nivel A2, aunque no se registró una diferencia significativa. También, dichos hallazgos permiten responder al objetivo general 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordadas con los niveles de competencia del ELE A2 y B1.

Formaron parte de los procedimientos definidos en este segundo estudio, la comparación entre los resultados del sub-corpus CAELE y los de un corpus de referencia, esto fue, el sub-corpus de hablantes nativos de español. Esta comparación se realizó con la finalidad de observar similitudes o diferencias en los usos colocacionales.

A continuación, se describen los resultados sobre el análisis colocacional de las producciones escritas por los nativos. Luego de la descripción, se comparan los resultados colocacionales de los niveles A2 y B1 con los resultados del sub-corpus de español utilizando las pruebas ANOVA.

8.3. Colocaciones en el sub-corpus de nativos de español

En este apartado se muestran las frecuencias obtenidas a partir del análisis de 83 tareas de escritura. El proceso de etiquetado fue realizado de igual forma con la taxonomía implementada en el *software UAM Corpus Tool*.

8.3.1. Colocaciones léxicas verbo-nominales

En el Gráfico 14 se muestra un total de 81 colocaciones léxicas producidas por hablantes nativos de español. 49 combinaciones correspondieron a verbos con significado pleno (60%) y 32, a verbos soporte (40%). Llama la atención que no existió una distancia colocacional abismante entre una y otra categoría, así como se observó en el sub-corpus de ELE.

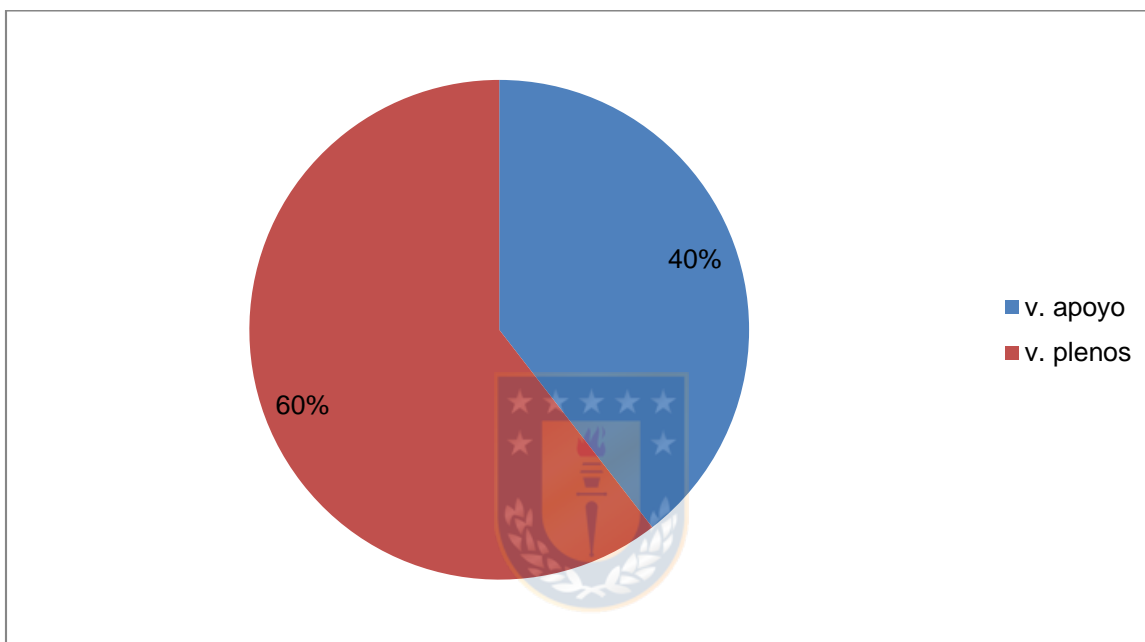


Gráfico 14. Colocaciones léxicas V+S en el sub-corpus de nativos. Elaboración propia.

8.3.2. Colocaciones léxicas verbo-nominales con verbos de apoyo

El análisis del corpus de nativos arrojó un total de 32 colocaciones léxicas simples con verbos des-lexicalizados (*ver Tabla 24*). Destacó el verbo “hacer”, cuyo valor léxico-semántico se desplazó y proyectó en el sustantivo que lo acompañaba. Por ejemplo, “hacer un llamado” (40, 63%), “hacer diferencia” (9, 38%) o “hacer ingreso” (9, 38%).

N°	Colocaciones	Frecuencia	%
1	hacer un llamado	13	40,63%
2	tomar una decisión	6	18,75%
3	tener conocimiento	4	12,50%
4	hacer diferencia	3	9,38%
5	hacer ingreso	3	9,38%
6	tomar conciencia	3	9,38%
Total	-	32	100%

Tabla 24. Colocaciones léxicas simples con verbos de apoyo en hablantes nativos de español. Elaboración propia.

En cuanto al contexto de uso de las colocaciones producidas por los hablantes nativos, se mencionan los siguientes ejemplos:

- 1) “Ademas la ex presidenta Bachelet hace un llamado a no retroceder en los avances que tuvo su gobierno en relación a la equidad de género” (hablante 2018S33T3CHN),
- 2) “Si cada vez hay más basura en el mar es porque cada vez generamos más, tenemos que empezar a tomar conciencia del daño que provocamos a otros y a nosotros mismos” (hablante 2018S20T4CHN) y
- 3) “Un sujeto que tiene conocimientos sobre sí mismo puede desenvolverse de mejor forma con sus compañeros de trabajo” (hablante 2018S12T4CHN).

8.3.3. Colocaciones léxicas con verbos de significado pleno

En la Tabla 25 se muestra las colocaciones con verbos de significado pleno obtenidas de las tareas producidas por los hablantes nativos. Destacó la colocación “dar cuenta” (32, 65%) con el valor léxico de comprender y “llamar la

atención” (18, 37%) con el sentido de advertir. Una colocación poco usual fue “sufrir violencia” (8, 16%). Estos resultados son novedosos, puesto que se esperaba un mayor número de colocaciones de este tipo por el hecho de que estas estructuras están conformadas por verbos con un significado léxico autónomo.

N°	Colocaciones	Frecuencia	%
1	dar cuenta	16	32,65%
2	dar oportunidad	3	6,12%
3	leer libro	4	8,16%
4	llamar la atención	9	18,37%
5	alcanzar igualdad	3	6,12%
6	brindar apoyo	3	6,12%
7	lanzar un libro	3	6,12%
8	lograr avances	4	8,16%
9	sufrir violencia	4	8,16%
Total	-	49	100%

Tabla 25. Colocaciones léxicas simples con verbos de significado pleno en hablantes nativos de español. Elaboración propia.

En este sentido, es relevante revisar algunos contextos de uso de las combinaciones con verbos lexicalizados:

- 1) *“Actualmente aunque exista la Ley Zamudio en el país que tiene por fin repeler todo acto de discriminación, estos actos continúan existiendo, con esto podemos dar cuenta de la realidad que existe”* (2018S21T3CHN),
- 2) *“Al leer libros de historia, matemática, biología, etc. Nos ayuda a tener conocimiento y no quedar como ignorante ante alguna charla”* (2018S33T5CHN) y
- 3) *“Se han logrado avances desde distintas plataformas y perspectivas, pero todavía queda bastante trabajo por hacer”* (2018S29T3CHN) y

- 4) “El deporte me llama demasiado la atención ya que es algo que me llena y me completa como persona” (2018S27T5CHN).

8.3.4. Bases de las colocaciones verbo-nominales con verbos de apoyo en el sub-corpus de nativos

En el Gráfico 15 se da cuenta de los sustantivos que fueron seleccionados por los nativos y que se combinaron con los verbos para así formar una colocación. El término “llamado” fue la base más frecuente (41%) en este sub-corpus lingüístico, seguido de “decisión” (19%) y conocimiento (13%). Se puede observar que todas estas clases léxicas pueden ser clasificadas como sustantivos abstractos al igual que las producidas en el nivel B1, por lo tanto, pareciera entonces que sí existe una relación entre este tipo de sustantivos y los verbos de apoyo, tal cual se observó en el sub-corpus de aprendientes de ELE.

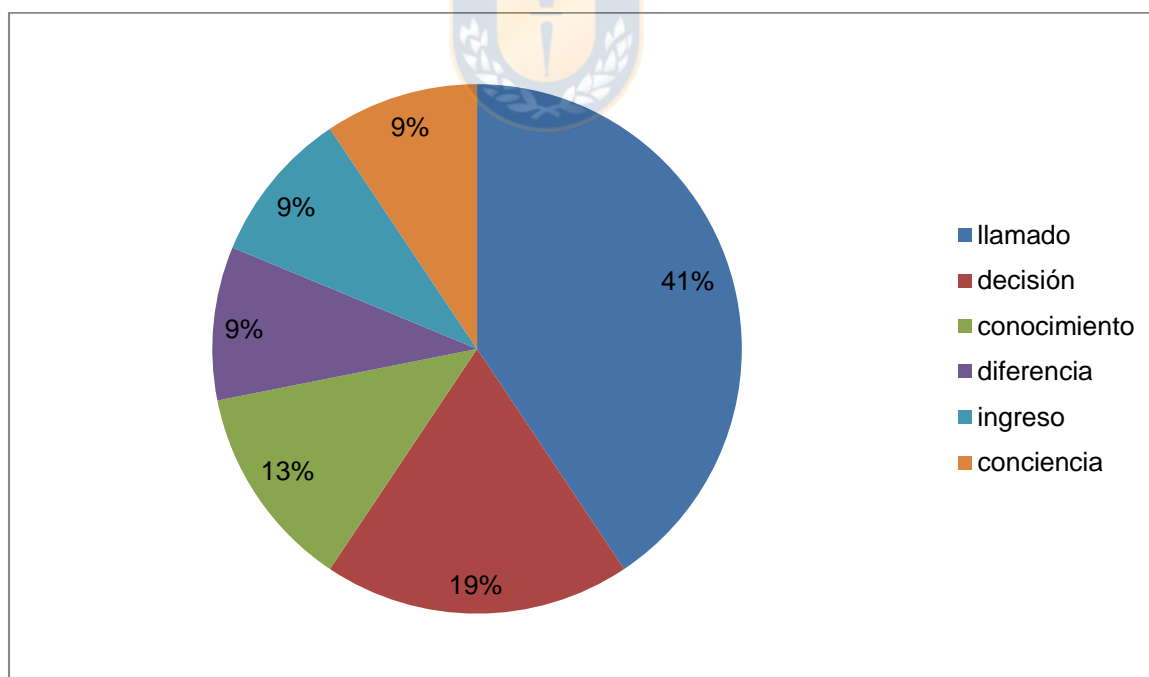


Gráfico 15. Bases con verbos de apoyo en el sub- corpus de nativos. Elaboración propia.

8.3.5. Bases de las colocaciones verbo-nominales con verbos de significado pleno en el sub-corpus de nativos

En el Gráfico 16 se da cuenta de las bases más producidas por los hablantes nativos. Se logra observar que solo hay 8. La mayoría de estos sustantivos se caracterizaron por tener una naturaleza abstracta. En cuanto al número gramatical se utilizó tanto el singular como el plural. La base más productiva fue “cuenta” (33%), “atención” (19%) y “libro” (14%). Por ejemplo,

“Finalmente y en el marco de la equidad de género recientemente fue lanzado el libro #Las niñas pueden” (2018S7T3CHN).

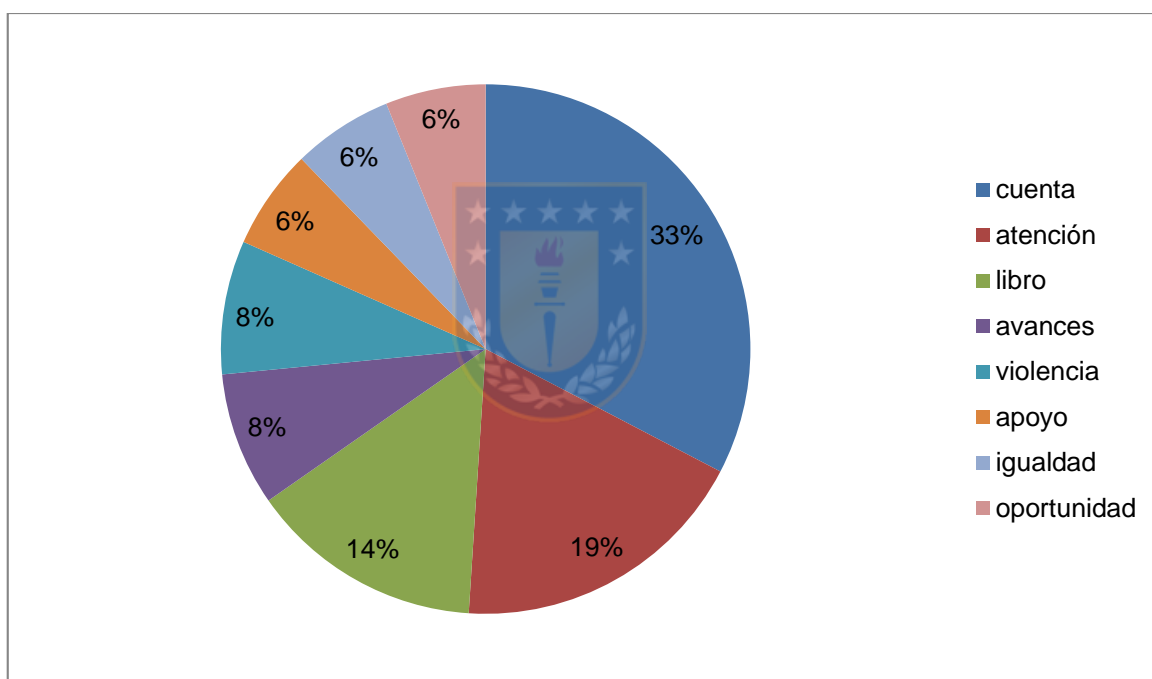


Gráfico 16. Bases con verbos plenos en el sub- corpus de nativos. Elaboración propia.

Se puede afirmar que las bases (“violencia”, “igualdad” y “oportunidad”) estuvieron relacionadas con el tema “Equidad de género”. En el caso de las demás bases, estas se correspondieron con dos temas en los cuales los hablantes nativos debieron escribir dos composiciones libres.

8.3.6. Colocados de las colocaciones verbo-nominales con verbos de apoyo en el sub-corpus de nativos

En el Gráfico 17, se observa la predominancia del verbo “hacer” (59%) en relación a los colocados “tomar” (28%) y “tener” (13%). Posiblemente se deba a que la función de esta clase léxica se encuentra consolidada en el lexicón mental de los hablantes nativos del español, por lo tanto, funciona como un verbo multifuncional con la capacidad de ser atraído por los sustantivos, bases de la colocación.

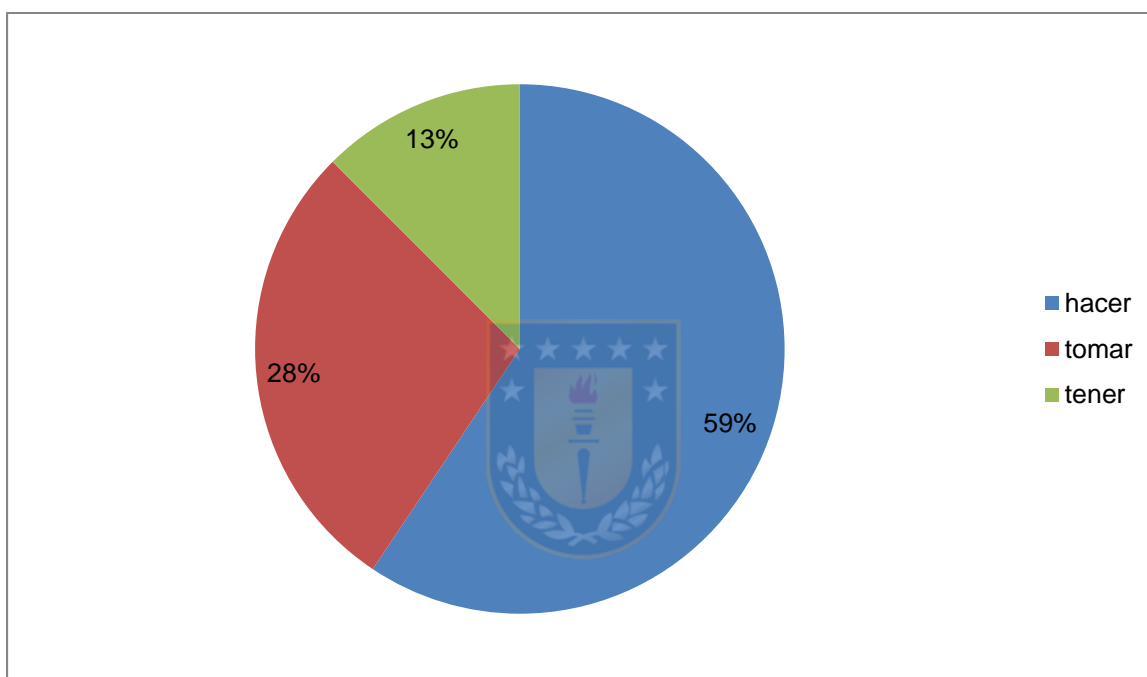


Gráfico 17. Colocados con verbos de apoyo en el sub-corpus de nativos. Elaboración propia.

8.3.7. Colocados de las colocaciones verbo-nominales con verbos de significado pleno en el sub-corpus de nativos

Los colocativos o colocados que se observan en el sub-corpus de nativos se distribuyeron en 7 verbos, tal como se muestra en el Gráfico 18. El colocado más frecuente fue “dar” (42%) y “llamar (20%)”, mientras que el menos productivo fue “lanzar” (7%).

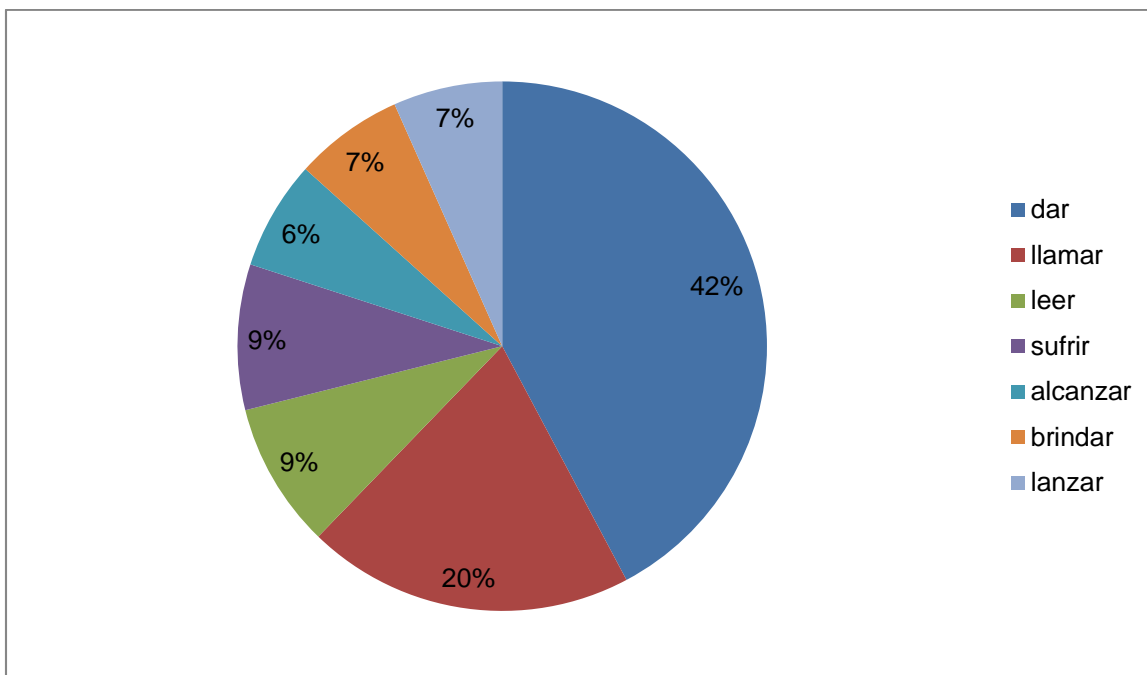


Gráfico 18. Colocados con verbos plenos en el sub-corpus de nativos. Elaboración propia.



8.4. Colocaciones léxicas V+S en el sub-corpus CAELE y el corpus de nativos

En el Gráfico 19 se compara la cantidad de colocaciones encontradas en el sub-corpus del corpus CAELE y en el sub-corpus de nativos de español. Los aprendientes de ELE produjeron una cantidad de 513 colocaciones léxicas verbo-nominales, mientras que los hablantes nativos solo, 81. En cuanto a las colocaciones léxicas simples conformadas con verbos plenos, 49 fueron identificadas en el sub-corpus de nativos, 248 en el A2 y 216 en el B1. Finalmente, en relación a los verbos de apoyo, se reportaron 32 en las tareas de escritura, pertenecientes a los nativos; 22, en las de nivel A2 y 27, en las de nivel B1. En esta categoría se puede apreciar hallazgos similares de estas estructuras, a diferencia de los verbos de contenido, en los cuales hay mayor diferencia numérica y también una amplia variación léxica en la producción colocacional entre los nativos de español y los aprendientes de ELE.

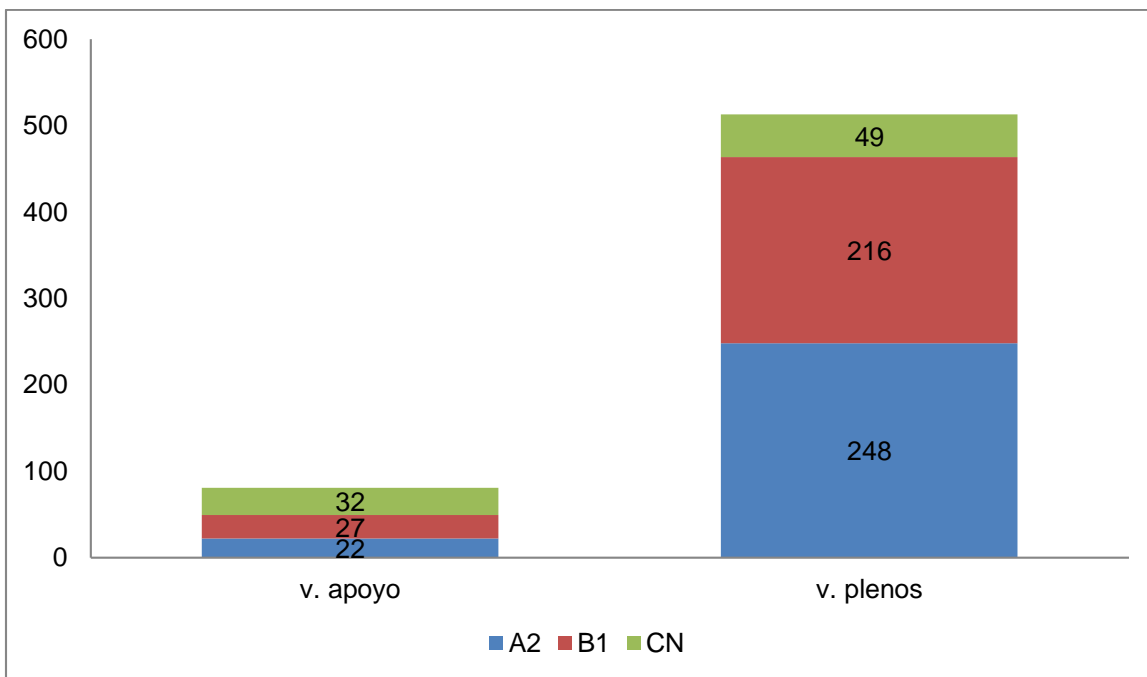


Gráfico 19. Comparación de colocaciones léxicas entre el sub-corpus CAELE y el sub-corpus de nativos. Elaboración propia.

8.5. Pruebas estadísticas ANOVA

A continuación, se presentan los resultados de la comparación de las colocaciones léxicas simples verbo-nominales en los dos niveles de competencia lingüística de ELE a la luz de los obtenidos en el sub-corpus de hablantes nativos de español. Los hallazgos se obtuvieron mediante la aplicación de las pruebas estadísticas ANOVA con la finalidad de determinar la varianza en la cantidad de colocaciones léxicas simples verbo-nominales (colocaciones léxicas simples con verbos de apoyo y colocaciones léxicas simples con verbos de significado pleno) producidas tanto por los dos grupos de aprendientes de ELE (A2 y B1) como por el grupo de hablantes nativos.

8.5.1. Colocaciones léxicas simples verbo-nominales con verbos de apoyo

Las siguientes tablas tienen por objetivo dar a conocer los hallazgos de la aplicación de las pruebas ANOVA.

En la Tabla 26, se observa que los hablantes de español produjeron un promedio mayor de estas estructuras (2,2857) que los grupos de aprendientes de ELE. La desviación estándar de las colocaciones fue más baja en el nivel A2 (2,0649), esto quiere decir, que hubo menor dispersión en la cantidad de colocaciones con verbos de apoyo producidas, a diferencia del nivel B1 y de los nativos de español.

Grupos	A2	B1	HN
Valor de P	P= 0,8495		
Promedio	1,5714	1,9286	2,2857
Desviación estándar	2,0649	3,8723	3,6675
Desviación de error	0,5519	1,0349	0,9802

Tabla 26. Resumen de los datos de colocaciones léxicas-simples con verbos de apoyo. Elaboración propia.

En la Tabla 27, los resultados de la aplicación de la prueba ANOVA arrojaron un área F= 0,1638, lo que tuvo como consecuencia un alfa de P= 0,8. En síntesis, no hubo diferencias entre los tres grupos, más bien similitud en el uso de colocaciones léxicas con verbos de apoyo.

Descriptor	Grado de libertad	Suma de los cuadrados	Media del cuadrado	Valor crítico de F	Valor de P
Entre grupos	2	3,5716	1,7858	0,1638	0,8495
Dentro de grupos	39	425,218	10,903		
Total	41	428,7896			

Tabla 27. Resumen de ANOVA de colocaciones léxicas-simples con verbos de apoyo. Elaboración propia.

8.5.2. Colocaciones léxicas simples verbo-nominales con verbos de significado pleno

En la Tabla 28, se presenta las estadísticas descriptivas de los tres grupos (A2, B2 y grupo de hablantes nativos) en torno a los 61 tipos de colocaciones léxicas simples con verbos de significado pleno, los cuales presentan distintas frecuencias. El promedio en A2 fue de 4,0656, en B1 fue de 3,541, mientras que en los hablantes nativos alcanzó 0,8033. Se observa entonces que hubo una diferencia significativa entre el promedio de los hablantes nativos (HN) y el de nivel A2, con la única diferencia que este último grupo fue menos homogéneo y tendió a una dispersión con una desviación estándar de 5,3.

Grupos	A2	B1	HN
Valor de P	P= 0,00024		
Promedio	4,0656	3,541	0,8033
Desviación estándar	5,363	5,3931	2,522
Desviación de error	0,6867	0,6905	0,3229

Tabla 28. Resumen de los datos de colocaciones léxicas-simples con verbos de significado pleno. Elaboración propia.

En cuanto a los estadígrafos de la prueba ANOVA sobre los resultados de colocaciones léxicas simples verbo-nominales con verbos de significado pleno, estos reportaron un área F de 8,7465. Dicho valor se tradujo en un alfa de $P=0,0002$, que fue altamente significativo (*ver Tabla 29*). En consecuencia, existió una diferencia estadísticamente significativa en cuanto al bajo índice de uso de verbos de plenos producidos por el grupo de hablantes nativos versus un índice mayor en los grupos A2 y B1.

Descriptor	Grado de libertad	Suma de los cuadrados	Media del cuadrado	Valor critico de F	Valor de P
Entre grupos	2	374,3938	187,1969	8,7465	0,0002
Dentro de grupos	180	3852, 4668	21,4026		
Total	182	4226,8606			

Tabla 29. Resumen de ANOVA de colocaciones léxicas-simples con verbos de apoyo. Elaboración propia.

Los resultados de la prueba ANOVA tienen gran impacto, porque evidencian en las colocaciones léxicas simples con verbos de significado pleno que el uso de estas combinaciones no es homogéneo entre los aprendientes del ELE y los hablantes nativos. En sentido, la comparación entre corpus lingüísticos (Jarvis, y Pavlenko, 2007) ratificada con la aplicación de pruebas estadísticas es fundamental para el estudio de las problemáticas lingüísticas en el contexto de enseñanza-aprendizaje del ELE.

En esta primera parte, se expusieron detalladamente los principales resultados de los dos estudios que constituyen esta tesis doctoral: 1) Frecuencia colocacional en el sub-corpus CAELE y 2) Comparación de la producción colocacional entre los niveles de competencia lingüística. De esta forma, esta sección respondió a los objetivos generales y objetivos específicos planteados en la metodología de investigación y también a la H1 y H2 postuladas.

En la segunda parte del capítulo, se discuten los resultados ya expuestos a la luz de literatura científica sobre la producción de colocaciones léxicas verbo-nominales en el ámbito de LM y ELE.

9. Discusión

El análisis sobre colocaciones léxicas simples verbo-nominales realizado en las tareas de escritura producidas por los aprendientes de ELE, pertenecientes a los dos niveles de competencia comunicativa ha sido relevante por la cantidad y los tipos de colocaciones producidas. Al considerar los tipos de colocaciones léxicas simples verbo-nominales junto a su frecuencia, estos datos aportan y otorgan mayor exhaustividad a la lista de nociones expuestas en el Plan curricular del Instituto Cervantes (2006).

En el contexto de enseñanza de ELE, los resultados evidencian como elemental el trabajo con este tipo de colocaciones léxicas desde el nivel de competencia inicial, dado que en el nivel A2, los aprendientes producen variadas combinaciones. Tal como lo plantea Pérez (2014) existe la necesidad de prestar atención a este tipo de unidades léxicas desde los primeros niveles. En este sentido, Navajas (2006) postula que, las colocaciones no solo se pueden trabajar en niveles intermedios y avanzados, sino también en los niveles más básicos.

Esta discusión se organiza por los objetivos generales planteados en la investigación y responde a sus respectivos objetivos específicos.

9.1. Objetivo general 1 “Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE”

En la actualidad, el Marco Común Europeo de Referencia para las Lenguas (Consejo de Europa, 2001) no cuenta con una lista explícita de colocaciones léxicas para los niveles de competencia lingüística del español descritos en su Plan Curricular; no obstante, contiene una lista³⁰ de exponentes, conocidos como nociones generales y nociones específicas. Estos referentes constituyen expresiones lingüísticas asociadas a nociones (significados) y a funciones

³⁰ Esta lista no contiene frecuencias, por lo tanto, la discusión en esta parte solo se enfoca en aspectos de carácter cualitativo.

(intenciones comunicativas) en consonancia con el enfoque nocio-funcional (Instituto Cervantes, 2006). Por ello, se ha considerado su repertorio para discutir los resultados de esta investigación con sus nociones, hasta el momento expresiones afines con las colocaciones. Inclusive, es necesario mencionar que un grupo reducido de las nociones del Plan Curricular conforma colocaciones léxicas verbo-nominales, por ejemplo: “hacer deporte”, “cancelar un contrato”, “tener conocimiento”, “perder las llaves”, “tener una duda”, “obtener un resultado”, “doler la cabeza”, entre otras. Pérez (2014) enfatiza que “la combinatoria ofrecida está lejos de ser exhaustiva y necesita ser completada para ofrecer así una referencia sobre las colocaciones que deberían formar parte de cada nivel” (p.19) en el ámbito de ELE. Siguiendo este planteamiento, los resultados arrojados en esta investigación contribuyen de forma explícita en la descripción de colocaciones verbo-nominales de niveles A2 y B1 con la finalidad de robustecer la lista ofrecida por el Plan Curricular del Instituto Cervantes (2006).

En relación al objetivo específico 1 “Identificar las colocaciones léxicas simples verbo-nominales de competencia del ELE A2 y B1 utilizando los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico”, se puede señalar lo siguiente:

- a) Por una parte, en relación a la revisión realizada sobre ambas categorías de nociones, se puede señalar que en el nivel A2 las colocaciones léxicas verbo-nominales con verbos plenos identificadas en este estudio tienen similitud con los exponentes específicos del MCER (Consejo de Europa, 2001) para las Lenguas. Por ejemplo, se encontró en el análisis la base “productos”, la cual apareció en la noción “actividad laboral”; también, la colocación “ganar dinero”, que tuvo registro en la noción “derechos y obligaciones laborales”, “tener tiempo” con registro en “tiempo libre y entretenimiento”. En las colocaciones con verbo de apoyo, el colocado “hacer” también apareció en distintas dimensiones generales debido a que es un verbo des-lexicalizado, por ejemplo, en las nociones generales “existencia e inexistencia” y “movimiento y estabilidad”. La colocación “tener

tiempo” tuvo correspondencia con la noción temporal “referencias generales”, la base “música”, que fue muy frecuente en los resultados de esta investigación, apareció en la noción “capacidad y competencia”. Se puede afirmar entonces que hay una relación de correspondencia entre este tipo de colocaciones, sus componentes, y las nociones planteadas en el Plan curricular del Instituto Cervantes (2006), las cuales a su vez se encuentran articuladas con el MCER (Consejo de Europa, 2001) para las Lenguas.

- b) Por otra parte, en el nivel de competencia B1, las bases “música” y “canciones” aparecieron en la noción “audibilidad y audición”, el colocado “ganar”, en la noción “éxito y logro”. La base “historias” y el colocado “dar” tuvieron registro en la noción “expresión verbal”, el colocado “hacer” apareció en “celebraciones y actos familiares, sociales y religiosos” y en “lenguaje de aula”. Finalmente, la colocación “contar la historia” tuvo correspondencia en la noción específica “la literatura”. En consecuencia y sobre los fundamentos de estas nociones, se puede observar que las colocaciones léxicas simples verbo-nominales producidas por los aprendientes de este nivel también mantienen una relación de correspondencia con las nociones expuestas en el Plan Curricular del Instituto Cervantes (2006).

En relación con el objetivo específico 2 “Clasificar las colocaciones léxicas simples verbo-nominales en construcciones con verbos de apoyo y en construcciones con significado pleno”, se puede declarar que:

- a) El uso de colocaciones con verbos de apoyo fue menor en relación a la producción de colocaciones con verbos de significado pleno. Esto se encuentra en consonancia con la prueba *t-test* realizada y con lo postulado por la literatura, dado que los verbos de apoyo constituyen categorías léxicas finitas y limitadas; a diferencia de los verbos de significado pleno, que cumplen la función de portar variada información semántica. Las

colocaciones con verbos de apoyo más frecuentes resultaron ser las que estaban formadas por el colocado “tener”. Estos resultados dialogan con los hallazgos encontrados por De Miguel (2011), aunque en su lista de verbos de apoyo también apareció como frecuente el verbo “echar”, el cual no registró ningún uso como colación en el sub- corpus CAELE. También los resultados del Corpus de Español³¹ (Davies, 2012) posicionan el verbo “tener” como frecuente en colocaciones verbo-nominales.

- b) Desde una perspectiva de los constituyentes gramaticales, las colocaciones verbo-nominales encontradas en el sub- corpus CAELE se caracterizaron por la presencia de modificadores, tales elementos correspondieron a artículos determinados e indeterminados. En el estudio dirigido Skorepova (2008) sobre colocaciones verbo-nominales producidas por aprendientes checos, las combinaciones identificadas también se caracterizaron por admitir elementos adicionales, tales como cuantificadores. Entonces, se puede señalar que si bien estas formas lingüísticas presentan mutua atracción entre la base y el colocado; esta atracción también incluye todos sus componentes: artículos, cuantificadores y adjetivos. Hasta aquí, con los resultados obtenidos del análisis del sub-corpus CAELE y los del sub-corpus de nativos de español se puede postular en términos de contribución que las colocaciones verbo-nominales no solo admiten la presencia de un verbo y de un sustantivo, sino que también la de otras categorías léxicas cuya función es modificar y precisar el significado léxico de la colocación.

En relación con el objetivo específico 3 “Delimitar las frecuencias de bases y colocados de las estructuras en estudio”.

- a) Las bases de las colocaciones estudiadas fueron definidas en su mayoría como sustantivos de naturaleza concreta, vinculadas a los tópicos de escritura. Las más frecuentes fueron “música” (139), “pintura” (34) y “artesanía” (32), pertenecientes a la categoría de colocaciones de verbos

³¹ Corpus de Español compuesto por más de un millón de palabras procedentes de más de 20.000 textos.

con significado pleno; mientras que, en la categoría de verbos de apoyo, las bases más productivas fueron “importancia” (24) e “influencia” (7), definidas como sustantivos abstractos. Al respecto no se ha encontrado información sobre estos sustantivos en su calidad de bases o nodos; no obstante, el Corpus de Referencia del Español Actual (CREA) afirma que la clase léxica “música” tuvo una aparición de 28928 casos en un total de 6889 documentos, por consiguiente, se puede señalar que sí constituye un ítem frecuente. “Pintura”, tuvo una aparición de 10697 casos distribuidos en 3008 documentos y “artesanía”, solo apareció 791 casos en 458 documentos. “Importancia” se constató 23761 veces en 9415 documentos y finalmente “influencia”, 10972 casos en 4226 documentos. En virtud de lo expuesto, se postula que estas bases tienen alta colocabilidad, es decir, son capaces de combinarse, seleccionar otras categorías léxicas y generar una colocación.

- b) Se presentó como tendencia de uso que los colocados de las colocaciones verbo-nominales constituidas por verbos de apoyo se conformaban por bases de carácter abstracto que seleccionaban verbos como “tener” y “dar”. Dicha tendencia ha sido poco explorada. Esta se logró corroborar con el colocado “hacer” a partir de los resultados del estudio del sub-corpus de español (Kloss, 2018). En estas tareas de escritura, las bases que seleccionaron los hablantes nativos eran de esta naturaleza y se combinaban con el colocado “hacer”. Para Molero y Salazar (2013) en los resultados de su investigación, las colocaciones con el verbo “tener” tienden a extender su significado de posesión a un significado más abstracto y metafórico. En sus hallazgos se observó que “tener” tiende a combinarse con sustantivos que se refieren a estados y cualidades. En este mismo sentido, para Skorepova (2008) el colocativo “tener” es considerado como un verbo frecuente en las colocaciones verbo-sustantivo en tareas de escritura producidas por aprendientes checos en el ámbito de ELE. En el Corpus de Español de (Davies, 2012) “tener” se posicionó como un

colocado más frecuente que “hacer”. A partir de estos resultados y en términos de contribución, se plantea el carácter abstracto de las bases que seleccionan colocados de verbos soporte para formar una colocación verbo-nominal.

9.2. Objetivo general 2 “Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1”

Por una parte, la tendencia de uso apuntó a que las colocaciones verbo-nominales encontradas se presentaron mayormente en el nivel inicial. Al parecer el desarrollo de la competencia colocacional comienza en los primeros niveles. Sin embargo, llama la atención que en el nivel B1 hayan disminuido en cuanto al número. Una razón podría ser que los aprendientes incorporan a sus producciones otro tipo de colocaciones léxicas tanto simples como complejas (Corpas, 1996; Koike 2001). Algunos ejemplos son: “sustantivo+ adjetivo”/ “adjetivo+ sustantivo” (“discusión acalorada”/“santa paciencia”), “verbo+ sustantivo” (“implorar misericordia”), “verbo+ preposición +sustantivo” (“intervenir en un conflicto”), “sustantivo+ preposición+ sustantivo” (“banco de peces”), “adverbio+ adjetivo/ participio” (“firmemente convencido”), “verbo+ adjetivo” (“salir mal parado”), “verbo + adverbio” (“cerrar herméticamente”).

Por otra parte, no se constató con los resultados del *t-test* una diferencia significativa en la producción de colocaciones verbo-nominales entre ambos niveles de competencia lingüística del español. No obstante, en lo que respecta a las colocaciones producidas por los aprendientes de ELE, estas fueron mayores que las producidas por los hablantes nativos de español. Estos resultados sí generaron una diferencia significativa al comparar los tres grupos con la aplicación de la prueba ANOVA. Además, la divergencia no fue solo numérica, sino también involucró la selección y variación léxica de los componentes.

Se esperaba que los nativos produjeran una mayor cantidad de colocaciones. Ya Granger (1998b) demostró que los aprendientes de lengua extranjeras usaban

muchas menos colocaciones que los hablantes nativos; sin embargo, la tendencia de la producción colocacional se manifestó de forma inversa en este estudio. Es interesante precisar que los resultados obtenidos son comparables con los obtenidos por Prieto, Mosqueira y Vásquez (2009). En su investigación titulada “Córpora y enseñanza de lenguas: se buscan colocaciones”, los investigadores realizaron un estudio comparativo en el que afirmaron que los resultados fueron sorprendentes, pues el número de colocaciones empleadas por estudiantes de español superó al utilizado por los hispanohablantes.

Una razón que intenta explicar la diferencia significativa entre ambos corpus lingüísticos apunta a que los contextos de producción fueron diferentes en las tareas de escritura. En un contexto con temáticas periodísticas, que involucró tres géneros (ensayo, columna de opinión y motivacional) y dos estilos (argumentativo y expresivo) producidos por los hablantes nativos de español, las colocaciones disminuyeron; mientras que, en un contexto con temáticas interculturales, que involucró solo el género informativo y el estilo narrativo/descriptivo en todas las tareas, las colocaciones aumentaron. Al respecto, para Pejovic (2010) la producción de colocaciones verbales estaría determinada por el aspecto contextual, es decir, por la situación comunicativa. Además, enfatiza que:

El contexto es uno de los factores imprescindibles para la comprensión de las palabras y de las combinaciones que forman. Adaptarse a la situación comunicativa y al tema que se trate también es de importancia vital para la producción adecuada y apropiada de enunciados, porque el contexto determina mucho la configuración del discurso (p.136).

Entonces, al parecer algunas colocaciones se pueden producir en ciertos contextos, los cuales favorecen su codificación y producción. Finalmente, el aprendiente ha de acabar intuyendo qué contextos permiten la codificación de un determinado significado por medio de un cierto significante. En consecuencia, se torna complejo corresponder en uno de los dos niveles de competencia comunicativa estudiados las colocaciones léxicas producidas por los hablantes nativos.

En relación al objetivo específico 1 “Comparar las colocaciones léxicas simples verbo-nominales con verbos de apoyo de acuerdo con los niveles del ELE A2 y B1”.

En términos de cantidad, las combinaciones con verbos de apoyo tuvieron mayor productividad en el nivel B1 que en el nivel A2, aunque la diferencia observada fue mínima. Ambos niveles de competencia comunicativa del español se caracterizaron por colocaciones léxicas con el colocado “tener”. En lo que respecta a la variación léxica de estas colocaciones y sus componentes hubo cierta similitud. Una razón que conviene reportar es el contexto temático en el cual fueron escritas las producciones y del cual ya se informó en el objetivo anterior. En esta misma línea, la importancia del contexto en el proceso de aprendizaje de idiomas se torna imprescindible.

Al consultar las nociones específicas del Plan Curricular del Instituto Cervantes (2006), se ha encontrado la colocación “tener tiempo” en las nociones del nivel de competencia A2, pero no hay hallazgos de colocaciones con verbos de apoyo. Entonces, la necesidad de incorporar una lista explícita de colocaciones verbo-nominales junto a sus frecuencias se hace evidente.

En relación al objetivo específico 2 “Comparar las colocaciones léxicas simples verbo-nominales con verbos plenos de acuerdo con los niveles del ELE A2 y B1”

Las colocaciones léxicas con verbos de significado pleno se presentaron con mayor productividad en el nivel de competencia lingüística A2. Sin embargo, en cuanto al tipo de colocaciones producidas estas fueron similares tanto en el nivel inicial como en el nivel intermedio y presentaron frecuencias parecidas. Las colocaciones “escuchar música” y “valorar la música” presentaron una frecuencia de 56 usos en el nivel inicial y en el nivel intermedio, una frecuencia de 53 usos. Al compararlas con las nociones específicas expuestas en el Plan Curricular del Instituto Cervantes

(2006), se puede establecer que no se encontró explícitamente ninguna de estas dos combinaciones. Estas se pudieron rastrear a través de su base “música”. En el nivel A2 este sustantivo apareció en la noción “capacidad y competencia” y en el nivel B1, en la noción “audibilidad y audición”. En este sentido, “música” puede ser definida como una base imprescindible en el proceso de enseñanza- y aprendizaje de ELE; por consiguiente, desde los primeros niveles de competencia lingüística, todo aprendiente debería comprender y producir colocaciones con esta base. Hasta el momento, existe la necesidad de robustecer las listas de frecuencias con colocaciones léxicas verbo-nominales en el ámbito de ELE que permitan establecer similitudes y diferencias a partir del estudio de distintas problemáticas lingüísticas. Por ejemplo, al ingresar la entrada música al Diccionario de Colocaciones del Español (DICE) esta no está contemplada.

En relación al objetivo específico 3 “Comparar las frecuencias de bases y colocados de las estructuras en estudio de acuerdo con los niveles del ELE A2 y B1”.

Tal cual se informó, la frecuencia de bases y colocados fueron similares en los niveles de competencia lingüística estudiados. En ambos niveles la base que presentó mayor productividad fue “música” y los colocados más frecuentes fueron “escuchar” y “valorar”. Prácticamente no se registraron colocaciones con el colocado “hacer” en ambos niveles de competencia del sub-corpus de ELE, mientras que en el sub-corpus de nativos, este tipo de combinación fue la más productiva. Uno de las colocaciones más utilizadas por los aprendientes de ELE en sus producciones escritas fue la que se conformaba por el colocado “tener”. García (2014) estimó en su análisis de combinaciones con verbos soporte que las colocaciones con verbo “tener” fueron más frecuentes en aprendientes de ELE que en los hablantes de español. Los nativos se caracterizaron por utilizar combinaciones que contenían el colocado “hacer”. Según Moncó (2013) es “el verbo menos marcado de acuerdo con los universales semánticos locativos propuestos por Traugott (1978)” (p.3). Es decir, es el menos marcado de los

transitivos, porque contiene un valor léxico muy poco restrictivo, hecho que provoca que en muchos casos se utilice como “verbo comodín” en reemplazo de otros verbos. Por ejemplo, “hacer un viaje” se puede reemplazar por el derivado semántico verbal “viajar” o “hacer un trabajo” por el derivado “trabajar”. Al respecto, estos hallazgos: a) dialogan positivamente con los resultados del estudio “Construcciones con verbos de apoyo. Caracterización en español y comparativa con el catalán”, dirigido por Romera (2017), quien afirma que en español “hacer” es un verbo de apoyo muy productivo. b), también se encuentran en la misma línea investigativa de Koike (2001), quien establece en su clasificación de los verbos soporte más productivos en español “dar” (“dar un pelotazo”), “tener” (“tener angustia”) y “hacer” (“hacer una reverencia”), c) dado su bajo contenido semántico, este tipo de clase léxica permite un gran número de combinaciones imposibles de recoger en los diccionarios de colocaciones, por lo que Bosque (2004, 2006) señala que los diccionarios no son listas cerradas y no pueden incluir todas las combinaciones posibles, d) el uso de este verbo también se podría explicar desde el conocimiento del contexto mental preexistente que tienen los hablantes nativos del español. Esto los conduce a utilizar variadas combinaciones con verbos soporte. Para Altenberg y Granger (2001):

los verbos (“tener” y “hacer”) son de muy alta frecuencia y presentan unas características muy interesantes desde el punto de vista contrastivo: 1) expresan sentidos básicos y predominan en una gran variedad de campos semánticos; 2) tienen equivalentes de alta frecuencia en la mayoría de idiomas; 3) son palabras polisémicas que tienen tanto usos generales, abstractos, dessemantizados o gramaticalizados como significados específicos, especializados o idiomáticos (p.174).

Se puede inferir que, debido a estas características, el procesamiento de este tipo de colocaciones es diferente del experimentado en las colocaciones con verbos pleno. Pérez (2017) propone que el proceso de adquisición de los colocativos

deslexicalizados es más lento y más dificultoso porque sus componentes (base y colocado) cuentan con un peso semántico distinto.

Respecto a la H1 postulada: “La producción de colocaciones léxicas simples con verbos de apoyo es menor que la producción con verbos de significado pleno”, los hallazgos de esta investigación y la aplicación del *t-test* permiten señalar que sí existe diferencia significativa en la producción de colocaciones entre estas dos categorías. La razón reportada fue que los verbos de apoyos constituyen categorías finitas a diferencia de los verbos plenos. Esto se proyecta también al ámbito de la producción colocacional. Por lo tanto, la selección léxica realizada por la base de la colocación se ve restringida en las producciones de ELE. Los resultados fueron ratificados también en las tareas de escritura producidas en el corpus de referencia de español de nativos (Kloss, 2018). De este modo, se afirma, aunque sin una intención concluyente, que los verbos plenos por su variedad léxica manifiestan la capacidad de ser potenciales colocados seleccionados por las bases, no así los verbos de apoyo.

Sin una intención concluyente, los hallazgos y los resultados obtenidos a partir de la aplicación de pruebas estadísticas permiten señalar en relación a la H2: “La producción de colocaciones léxicas simples verbo-nominales es mayor en el nivel de competencia lingüística B1 que en el nivel A2”, que: no existe diferencia significativa entre el nivel B1 y A2 de los aprendientes, aunque se presentó una tendencia que favorece la producción en el nivel inicial de competencia lingüística.

Además los resultados permiten sugerir incipientemente que el aprendizaje de colocaciones léxicas conformadas por verbo y sustantivo se produce en los primeros estadios de aprendizaje por su inteligibilidad; no obstante, a medida que los aprendientes progresan en el desarrollo de su competencia lingüística de ELE, las colocaciones de esta naturaleza disminuyen y probablemente se deba a la incorporación de otras unidades fraseológicas y colocaciones léxicas, caracterizadas por mayor complejidad léxica y sintáctica.

En consecuencia, dado que la hipótesis 2 se rechazó, se da lugar a la hipótesis nula “La producción de colocaciones léxicas simples verbo-nominales no es mayor en el nivel de competencia lingüística B1 que en el nivel A2”.

Conviene mencionar también la relevancia del contexto de producción, dado que estaría constituyendo una variable que podría facilitar o restringir el uso de las colocaciones estudiadas. La idea anterior podría llevar a afirmar que, a menor dispersión de estilos, temáticas y géneros discursivos, la producción colocacional aumenta. Por el contrario, entre mayor dispersión de estilos y géneros discursivos, la producción colocacional disminuye.

A la luz de los planteamientos expuestos en este capítulo, la pregunta de investigación ¿Varían las colocaciones léxicas simples con verbos de apoyo y con verbos de significado pleno (colocaciones verbo-nominales) producidas por los aprendientes de niveles del ELE A2 y B1? ha sido respondida con el análisis realizado, que involucró la identificación de las colocaciones objeto de estudio y la comparación colocacional entre ambos niveles utilizando pruebas estadísticas. Los resultados obtenidos han permitido responder la interrogante planteada. En este sentido, la cantidad de colocaciones léxicas verbo-nominales no varió significativamente entre uno y otro nivel; sin embargo, los hallazgos demuestran una tendencia que favoreció a los aprendientes del nivel inicial.

CAPÍTULO IV. CONCLUSIONES, LIMITACIONES Y PROYECCIONES DE ESTA INVESTIGACIÓN

10. Conclusiones

En esta tesis doctoral se han presentado los principales fundamentos teóricos, la metodología de investigación y los resultados del análisis efectuado en torno a las colocaciones léxicas verbo-nominales. La investigación se ha enmarcado disciplinarmente en los procedimientos de Corpus de Aprendientes de Lenguas en Formato Electrónico. Los objetivos planteados fueron 1) Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE y 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1.

Según el interés que suscita la Lingüística de Corpus en el proceso de enseñanza-aprendizaje de ELE a partir de la observación empírica de distintas formas lingüísticas (Cruz Piñol, 2012), se ha investigado la producción colocacional en el sub-corpus CAELE. El análisis colocacional efectuado ha considerado los roles del léxico. Basándose en este componente, los usos de las colocaciones léxicas verbo-nominales han constituido un insumo básico que interactúa en un conjunto de niveles de la lengua (sintaxis + morfología + lexicón) en el contexto de enseñanza-aprendizaje de ELE.

A modo de conclusión, sobre los resultados obtenidos y considerando los estudios realizados en torno a las colocaciones en ELE, los datos hasta aquí expuestos ratifican algunas de las tendencias encontradas en dichos estudios: 1) las colocaciones léxicas verbo-nominales con verbos de significado pleno son mayores que las con verbos de apoyo, 2) los hallazgos sugieren la identificación y la enseñanza explícita de colocaciones desde los primeros niveles de competencia comunicativa del español por parte del profesor de ELE (Fernández, 2014), porque según el Plan Curricular del Instituto Cervantes (2006) se debe integrar la enseñanza de los fraseologismos desde el primer nivel y 3) Las combinaciones

estudiadas en esta investigación se consideran como elementales para el alumno de ELE, ya que permiten mejorar su competencia lingüística y acercarse a la forma en que un hablante nativo se expresa (Uriel, 2014) con la finalidad de cumplir distintos objetivos comunicativos.

Finalmente, las colocaciones léxicas verbo-nominales estudiadas contribuyen a la descripción de la competencia comunicativa del ELE, particularmente, a la competencia colocacional de niveles iniciales e intermedios, puesto que han sido identificadas en un sub-corpus de aprendientes que refleja la lengua en uso. Este sub-corpus ha utilizado la elección de tareas de escritura con temáticas interculturales. Por lo tanto, las combinaciones observadas: 1) tienen aplicabilidad dentro del campo de la enseñanza y el aprendizaje del ELE a través de variadas metodologías y tareas, 2) benefician tanto a docentes como aprendientes de ELE, e 3) incrementan el listado de nociones específicas y nociones generales que componen el Plan curricular del Instituto Cervantes (2006).



11. Limitaciones y proyecciones de la investigación

Esta investigación presenta algunas limitaciones. En primer lugar, el problema terminológico que actualmente se suscita en torno a la definición de lo que es una colocación constituyó una limitación cuando se comenzó a resolver la pregunta de investigación. Para Osorio y Serra (2012) “uno de los términos más controvertidos de la lexicología contemporánea es el de colocación” (p.105). Aún la literatura científica no ha establecido un consenso en cuanto a todos los rasgos y los criterios que deben poseer las estructuras de esta naturaleza. Esto conllevó a que el proceso de identificación en cierta instancia metodológica fuera complejo. Si bien la frecuencia aporta un criterio valioso y reconocido desde la tradición proveniente de los estudios fraseológicos, no es el único. Su selección va a depender entre otras variables del propósito de la investigación y de los recursos tecnológicos que se utilicen.

En segundo lugar, los contextos de producción constituyeron una restricción (temáticas, géneros y estilos discursivos) que afectó la producción colocacional producida entre los aprendientes de ELE y los nativos de español. Esta limitación se constató en la fase de los resultados, puesto que existen escasos estudios que aborden la relación entre la producción colocacional en el ámbito de ELE y los contextos de producción.

Por último, el tamaño de la muestra que produjo los textos fue reducido (n=62 en ELE y n=35 en español), por lo tanto, los resultados de esta investigación no permiten extraer conclusiones definitivas y limitan su validez externa.

Como proyección de investigación de esta tesis doctoral, se encuentra dentro de las líneas futuras: a) proponer un listado representativo y elemental de colocaciones léxicas simples verbo-nominales considerando las nociones específicas y las nociones generales planteadas en el Plan Curricular del Instituto Cervantes (2006), b) generar un test que permita evaluar las colocaciones léxicas estudiadas acorde a los niveles de competencia lingüística A2 y B1, propuestos por el MCER para las Lenguas (Consejo de Europa, 2001) y que contribuya a describir la competencia colocacional desde el enfoque de la evaluación en ELE, c) asimismo, también surge la motivación y la necesidad de aumentar el tamaño de los participantes y las tareas de escritura incorporando nuevos niveles de proficiencia para robustecer el sub-corpus y obtener resultados con mayor representatividad.

12. Referencias bibliográficas

Aguilar-Amat, A. (1993). *Las colocaciones de nombre y adjetivo. Un paso hacia una teoría léxico-semántica de la traducción*. Barcelona, España: Universidad Autónoma de Barcelona.

Alexopoulou, A. (2005). Aproximación al tratamiento del error en la clase de E/LE desde la perspectiva del análisis de errores. *Estudios de Lingüística Aplicada*, 23(41), 101-125.

Alexopoulou, A. (2006). Los criterios descriptivo y etiológico en la clasificación de los errores del hablante no nativo: una nueva perspectiva. *Porta Linguarum*, 5,17-35.

Alexopoulou, A. (2007). Errores intralinguales e interlinguales en la interlengua escrita de aprendientes griegos de ELE. Trabajo presentado en el *Congreso de la Asociación Internacional de Hispanistas XVI: Nuevos caminos del hispanismo*. Madrid, España. Recuperado de https://cvc.cervantes.es/literatura/aih/pdf/16/aih_16_2_004.pdf

Alonso, M. (1993). *Las funciones léxicas en el modelo lexicográfico de I. Mel'čuk* (tesis doctoral). Universidad Nacional de Educación a Distancia, Madrid, España.

Alonso, M. (1994). Hacia una definición del concepto de colocación: De J. R. Firth a I. A. Mel'čuk. *Revista de Lexicografía*, 1, 9-28.

Alonso, M. (1997). Coocurrencia léxica y descripción lexicográfica del verbo DAR: hacia un tratamiento de los verbos soportes. *Zeitschrift für Romanische Philologie*, 113(3), 380-417.

Alonso, M. (2002). Colocaciones y contorno de la definición lexicográfica. *Revista LEA*, 24(1), 63-96.

- Alonso, M. (2004). *Las construcciones con verbos de apoyo*. Madrid, España: Visor Libros.
- Alonso, S. y Kalan, M. (2004). La necesidad de un diccionario combinatorio: La importancia de las colocaciones en la enseñanza de ELE. En *Las Gramáticas y los Diccionarios en la Enseñanza del Español como Segunda Lengua: Deseo y Realidad*. Ponencia llevada a cabo en el XV Congreso Internacional de la ASELE, Sevilla, España.
- Altenberg, B. y Granger, S. (2001). The grammatical and lexical patterning of “make” in native and non-native student writing. *Applied Linguistics*, 22(2), 173–195.
- Al-Zahrani, M. (1998). *Knowledge of English lexical collocations among male Saudi students majoring in English at a Saudi University* (tesis doctoral). University of Pennsylvania, Pensilvania. Indiana.
- Bally, Ch. (1951). *Traité de stylistique française*. París, France: Librairie Klincksieck.
- Baralo, M. (2007). *Adquisición de palabras: Redes semánticas y léxicas*. En *Aprender y enseñar léxico*. Ponencia llevada a cabo en el Foro de español internacional, organizado por la Universidad Antonio Nebrija, la editorial SGEL y el Instituto Cervantes. Múnich, Alemania.
- Beas, T. (2009). Transferencia léxica en las colocaciones con hacer y dar en el español de Mallorca desde una perspectiva diacrónica. In L. Romero Aguilera y C. Julia Luna (Eds.), *Tendencias actuales en la investigación diacrónica de la lengua: Actas del VIII Congreso Nacional de la Asociación de Jóvenes Investigadores de Historiografía e Historia de la Lengua Española*. Barcelona, España.
- Benson, M., E. Benson, R. y F. Ilson (1986). *Lexicographic Description of English*. Amsterdam, Holland: John Benjamins Publishing Company.

- Bestgen, Y. y Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach, *Journal of Second Language Writing*, (26), 28-41.
- Biber, D., Conrad, S. y Reppen, R. (1998). *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge, England: Cambridge University Press.
- Biskup, D. (1992). L1 influence on learners' renderings of English collocation: a Polish/German empirical study. En P. Arnaud y H. Bejoint (Eds.), *Vocabulary and Applied Linguistics* (pp.85-93). Londres, England: Macmillan,
- Blanco, X. (2000). Verbos soporte y clase de predicados en español. *LEA: Lingüística española actual* 22(1), 99-118
- Blasco, E. (2002). La lexicalización y las colocaciones. *Revista LEA*, (24), 35-61.
- Bosque, I. (2001). Sobre el concepto de colocación y sus límites, *Lingüística española actual*, 23(1), 9-41.
- Bosque, I. (2004). *Redes. Diccionario combinatorio del español contemporáneo*. Madrid, España: Ed. SM.
- Bosque, I. (2006). *Redes. Diccionario combinatorio del español contemporáneo*. Madrid, España: Ed. SM.
- Burger, H. (1973). *Idiomatik der Deutschen*. Tubinga, Germany: Max Niemeyer Verlag.
- Burger, H., Buhofer, A. y Sialm, A. (1986). *Handbuch der Phraseologie*. Berlín, Germany: Walter de Gruyter,
- Burgschmidt, E. y Perkins, C. (1985). *EB Fehlerkartei Englisch: Phraseologie: Kollokationen- Phraseme- Idiome*. Braunschweig, Germany: Privatepublisher.

- Butler, C. (1998). Multi-word lexical phenomena in functional grammar. *Revista Canaria de Estudios Ingleses*, (36), 13-36.
- Buyse, K. (2011). ¿Qué corpus utilizar para qué fines en la clase de ELE? En *Del texto a la lengua: la aplicación de los textos a la enseñanza-aprendizaje del español L2-LE*. Ponencia llevada a cabo en el XXI Congreso Internacional de ASELE, Salamanca, España.
- Carneado, Z. (1985). Notas sobre las variantes fraseológicas. *Anuario de Artes y Letras, Facultad de Artes y Letras, Universidad de Oriente*, (3), 62-72.
- Casares, J. (1950). Introducción a la lexicografía moderna. Madrid, España: C.S.I.C
- Castillo, C. (2013). Las construcciones con verbos de apoyo en un corpus comparable de textos jurídicos. *Edizioni Università di Trieste*, (2)93-106.
- Cid, J.P. (2012). *Tratamiento de las unidades pluriverbales en el Diccionario de uso del español de Chile: Propuesta de delimitación entre locuciones, colocaciones y compuestos sintagmáticos* (tesis de maestría). Universidad de Chile, Santiago, Chile.
- Collins COBUILD advanced learner's dictionary (2006). Londres, England: Harper Collins Publisher.
- Consejo De Europa (2001). Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación. Madrid, España: Instituto Cervantes-MECD, Anaya.
- Corder, P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5, 1-4.
- Corpas, G. (1996). Manual de fraseología española. Madrid, España: Gredos.

- Corpas, G. (2001a). Apuntes para el estudio de la colocación. *Revista LEA*, 23(1), 41-56.
- Corpas, G. (2001b). En torno al concepto de colocación. *Euskera*, (46), 89-108.
- CREA (1994). *Corpus de referencia del español actual (CREA)*. Real Academia Española. Disponible en <http://corpus.rae.es/creanet.html>
- Coseriu, E. (1977), Principios de semántica estructural, Madrid: Gredos. Cruse, D. A. (1986), Lexical Semantics, Cambridge: Cambridge University Press.
- Cruse, D. A. (1986). Lexical Semantics. Cambridge, England: Cambridge University Press.
- Cruz Piñol, M. (2012). *Lingüística de corpus y enseñanza del español como 2/L*. Madrid, España: Arco/ Libros S.L.
- Crystal, D. (1991). *The Cambridge Encyclopedia of Language*. Cambridge, UK: Cambridge University Press.
- Dagneaux, E., Denness, S. y Granger, S. (1998). Computer-aided error analysis. *System*, 26(2), 163-174.
- Davies, M. (2012). El corpus del Español. Recuperado de <https://www.corpusdelespanol.org/>
- De Alba Quiñones, V. (2014). El uso de las colocaciones en muestras escritas de aprendices de español. Análisis e implicaciones metodológicas. *Revista ELUA*, (28), 151-176. Doi: https://rua.ua.es/dspace/bitstream/10045/48504/1/ELUA_28_06.pdf
- De Alba Quiñones, V. (2009). El análisis de errores en el campo de ELE. Algunas cuestiones metodológicas. *Revista Nebrija de Lingüística aplicada a la enseñanza de Lenguas*, 5. Recuperado de <https://www.nebrija.com/revista->

linguistica/el-analisis-de-errores-en-el-campo-del-espa%C3%B1ol-como-lengua-extranjera.html

De Miguel, E. (2008). Construcciones con verbos de apoyo en español. De cómo entran los nombres en la órbita de los verbos. En Inés O., Manuel C. y Ramón G. (Eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)* (567-578). Pamplona, España: Servicio de Publicaciones de la Universidad de Navarra.

De Miguel, E. (2011). En qué consiste ser verbo de apoyo. En M.V. Escandell, M. Leonetti, y M. C. Sánchez, (Ed.), *60 problemas de gramática* (pp.139-146). Madrid, España: Akal editores.

DICE. (2004). Diccionario de Colocaciones del Español (DICE). Disponible en <http://www.dicesp.com/paginas>

Dobrovolskij, D. (1992). Phaseological Universals. Theoretical and Applied Aspects. En M. Kefer y J. Van Der Auwera (Eds.), *Meaning and Grammar. Cross- Linguistic Perspectives* (pp. 279-301). Berlin, Germany: Mouton de Gruyter.

Durrant, P. y Schmitt, N. (2009). To what extent do native and non-native speakers make use of collocations?, *IRAL* (47), 157-177.

EAGLES. (1996). *Synopsis and comparison of morphosyntactic phenomena encoded in lexicons and corpora. A common proposal and applications to European languages*. Pisa. Italy: ILC-CNR.

Elejalde, J. (2017). Errores de transferencia en textos escritos en Español como Lengua Extranjera producidos en ambientes de aprendizaje mediatizados por la tecnología (tesis de doctorado). Universidad de Concepción, Concepción, Chile.

- Ettinger, S. (1992). Unidades pluriverbales lexicalizadas y habitualizadas. En Haensch, G. et al. (Ed.), *La lexicografía. De la lingüística teórica a la lexicografía práctica* (pp.249-258). Madrid, España: Gredos.
- Evert, S. (2009). Corpora and collocations. En A. Lüdeling, y M. Kytö (Ed), *Corpus linguistics: An international handbook* (pp. 1212-1248). Berlin, Alemania: Mouton Gruyter.
- Farghal, M. y Obiedat, H. (1995). Collocations: A neglected variable in EFL. *IRAL*, 22(4), 315-331.
- Felix, S. (1981). The effect of formal instruction on second language acquisition. *Language Learning*, 31, 87-112.
- Fernández, M. (2005). La enseñanza de la fraseología. Evaluación de recursos y propuestas didácticas. En *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua, deseo y realidad*. Ponencia llevada a cabo en el XV Congreso Internacional de ASELE, Sevilla, España.
- Fernández, G. (2014). Enseñanza-aprendizaje de las colocaciones en nivel inicial (A1-A2). *Marcoele. Revista de Didáctica Español como Lengua Extranjera*, (19), 1-16.
- Fernando, C. y Flavell, R. (1981). *On idioms. Critical views and perspectives*. Exeter, England: University of Exeter Press.
- Ferrando, V. (2012). Aspectos teóricos y metodológicos para la compilación de un diccionario combinatorio destinado a estudiantes de E/LE (tesis doctoral). Universitat Rovira i Virgili, Tarragona, España.
- Ferreira, A. (2011-2014). *FONDECYT- CONICYT Nº 1110812*. "Un sistema tutorial inteligente para la focalización en la forma en la enseñanza del español como lengua extranjera". Proyecto de Investigación. Conicyt, Chile.

- Ferreira, A. (2014- 2017). *FONDECYT - CONICYT N°1140651*. “El feedback correctivo escrito directo e indirecto en la adquisición y aprendizaje del español como lengua extranjera”. Proyecto de Investigación. Conicyt, Chile.
- Ferreira, A. (2016). *Prueba de Multinivel con Fines Específicos Académicos*, Proyecto Fondecyt 1180974. Conicyt, Chile.
- Ferreira, A. (2014-2018). *Corpus de Aprendientes de Español Lengua Extranjera, CAELE*. Proyecto Fondecyt 1180974. Conicyt, Chile.
- Ferreira, A., Elejalde, J. y Vine, A. (2014). Análisis de Errores Asistido por Computador basado en un Corpus de Aprendientes de Español como Lengua Extranjera. *Revista Signos*, 47(86), 385-411.
- Ferreira, A., y Elejalde, J. (2017). Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE. *Revista Brasileira de Lingüística Aplicada*, 17(3), 509-538.
- Ferreira, A. (2018-2021). *FONDECYT- CONICYT N° 1180974*. “Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua”. Proyecto de Investigación. Conicyt, Chile.
- Ferreira, A., y Elejalde, J. (2019). Hacia un perfil lingüístico-comunicativo del estudiante de Español como Lengua Extranjera para fines Académicos. *Revista Nebrija De Lingüística Aplicada a La Enseñanza De Lenguas*, 13(27), 145-165. <https://doi.org/10.26378/rmlael1327326>
- Ferreira, A., y Elejalde, J. (2020). Propuesta de una taxonomía etiológica para etiquetar errores de interlengua en el contexto de un corpus escrito de aprendientes de ELE. *Forma y Función*, 33(1), 115-146. <https://doi.org/10.15446/fyf.v33n1.84182>

- Fleischer, G. (1982). *Phraseologie der deutschegegenwartssprache*. Leipzig, Germany: VEB.
- Firth, J. R. (1957). Modes of Meaning. En J.R. Firth (Ed.), *Papers in linguistics 1934-1951* (pp.190-215). Londres, England: Oxford University Press.
- Gairns, R., y Redman, S. (1986). *Working with Words*. Cambridge, England: Cambridge University Press.
- García, M. (2014). O uso de construcións con verbos soporte en aprendices de español como lingua estranxeira e en falantes nativos, *Cadernos de Fraseoloxía Galega*, (16), 181-198.
- García, M. (2017). Frecuencia y corrección colocacional en la producción escrita de aprendices de español. *Onomázein*, (39), 22-46.
- García, M. y Ramos, M. (2018). Asignación de niveles de aprendizaje a las colocaciones del Diccionario de Colocaciones del español. *Revista Signos*, 51(17), 153-174.
- García-Page, M. (2001). ¿Son las expresiones fijas expresiones fijas? *Revista Moenia*, (7), 165-197.
- González, F. (2006). Las colocaciones en la enseñanza del español de los negocios. *MarcoELE: Revista de Didáctica Español como Lengua Extranjera*, (2), 1-39.
- González, C. (2012). Skype y la interacción oral nativo/no nativo: funciones y rutinas conversacionales en Corinei, un corpus de interlengua español e italiano. In: Hernández, C.; Carrasco, A.; Álvarez, E. (Eds.). *La Red y sus*

aplicaciones en la enseñanza-aprendizaje del español como lengua extranjera. Valladolid: Universidad de Valladolid, pp. 283-293.

Granger, S. (1998a). *Learner English on computer*. Londres, England: Longman.

Granger, S. (1998b). Prefabricated patterns in advanced EFL writing: Collocations and formulae. En A.P. Cowie (Ed.) *Phraseology: Theory, analysis and applications* (pp.145-160). Oxford, England: Clarendon Press.

Granger, S. (2002). A bird's-eye view of learner corpus research. En S. Granger, J. Hungand y S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3-33). Philadelphia: EE.UU: John Benjamins.

Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, 20(3), 465-480.

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. *Language and Computers*, 52(1), 123-145.

Granger, S. (2012) How to use foreign and second language learner corpora. Mackey, Alison, y Gass, Susan M. (Eds.), *Research methods in Second Language Acquisition: A practical guide*, (pp. 7-29). London: Blackwell Publishing.

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal, *International Journal of Learner Corpus Research*, 1(1), 7-24. DOI: <https://doi.org/10.1075/ijlcr.1.1.01gra>

Granger, S. y Bestgen, Y. (2014). "The use of collocations by intermediate vs. advanced non-native writers: A bi-gram based study", *IRAL* 52(3), 229-252.

Haensch, G. (1985). *Los diccionarios del español en el umbral del siglo XXI*. Salamanca, España: Ediciones Universidad de Salamanca.

Halliday, M. (1961). Categories of the theory of grammar. *Word*, 17(3), 241-292.

- Halliday, M. (1966). "Lexis as a linguistic level". En C. Bazell, J. Catford, M. Halliday y R. Robins (Eds.), *In memory of John Firth* (pp.148-162). Londres, England: Longman.
- Hanks, P. (2000). "Dictionaries of idioms and phraseology en English". En G. Corpas Pastor (Ed.), *Las Lenguas de Europa: Estudios de Fraseología, Fraseografía y Traducción* (pp.303-320). Granada, España: Comares.
- Hasselgard, H. y Johansson, S. (2011). Learner corpora and contrastive interlanguage analysis. En F. Meunier, S. De Cock, G. Gilquin y M. Paquot (Eds.), *A taste for corpora: In honour to Sylviane Granger* (pp.33-62). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hausmann, F. J. (1985). Kollokation en imdeutschen Wörterbuch: ein Beitrag zur Theorie des lexikographischen Beispiels. *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch*". En H. Berrgenholtz y J. Mugdon (Eds.), *Lexicographica. Series Maior 3* (pp. 118-129). Tubinga, Germany: Max Niemeyer Verlag.
- Hausmann, F. (1979). Un dictionarie des collocation est-il possible? *Travaux de Linguistique ed Littérature*, 17(1), 187-195.
- Hernández, R., Fernández, C., y Baptista, P. (2014). *Metodología de la investigación*. México DF, México: Mc Graw Hill.
- Higueras, M. (1997). Las unidades léxicas y la enseñanza del léxico a extranjeros. *REALE*, 8, 35- 49.
- Higueras, M. (2004). Necesidad de un diccionario de colocaciones para aprendientes de ELE. En *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua, deseo y realidad*. Ponencia llevada a cabo en el XV Congreso Internacional de ASELE, Sevilla, España.

- Higueras, M. (2006). *Estudio de las colocaciones léxicas y su enseñanza en español como lengua extranjera*. Madrid, España: Ministerio de Educación y Ciencia.
- Hunston, S. (2006). Corpus Linguistics. En K. Brown (Ed.), *The encyclopedia of language and linguistics* (2a ed.) (pp. 234-248): Elsevier Science.
- Instituto Cervantes (1997-2019). *Diccionario de términos clave en ELE*. Centro Virtual Cervantes. Madrid: Instituto Cervantes- Biblioteca nueva.
- Instituto Cervantes (2006). *Plan curricular del Instituto Cervantes. Niveles de Referencia para el español*. Madrid, España: Instituto Cervantes- Biblioteca Nueva.
- Írsula, J. (1994). Entre el verbo y el sustantivo, ¿quién rige a quién? El verbo en las colocaciones sustantivo-verbales *Revista da Facultad de Letras- Línguas e Literatura*, 277-286.
- Izquierdo, M. (2004). La selección del léxico en la enseñanza del español como lengua extranjera. Su aplicación en el nivel elemental en estudiantes francófonos (tesis doctoral). Universitat de València, Valencia, España.
- Jarvis, S., y Pavlenko, A. (2007). *Crosslinguistic influence in language and cognition*. New York: Routledge.
- Kaszubski, P. (2000). Selected aspects of lexicon, phraseology and style in the writing of polish advanced learners of English: A contrastive, corpus- based approach (doctoral dissertation). *Adam Mickiewicz University, Poznań, Polonia*.
- Kellerman, E. (1978). "Transfer and non-transfer: Where are we now?" *Studies in Second Language Acquisition*, (2), 37-56.
- Kloss, S. (2018). *Los efectos del feedback correctivo escrito en el uso de marcadores discursivos con valor organizativo en crónicas periodísticas* (tesis de doctorado). Universidad de Concepción, Concepción, Chile.

- Kennedy, G. (1998). *An introduction to corpus linguistics*. London y NY: Longman.
- Koike, K. (2001). *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Madrid, España: Universidad de Alcalá y Universidad de Takushoku.
- Koike, K. (2002). Comportamientos semánticos en las colocaciones léxicas. *LEA. Lingüística Española Actual*, (24)1, 5-23.
- Koike, K. (2005). Colocaciones complejas en el español actual. En R. Pérez, E. Trives y G. Wotjak (Eds.), *Fraseología contrastiva: Con ejemplos tomados del alemán, español, francés e italiano* (pp. 169-184). Murcia, España: Univerdidad de Murcia.
- Krashen, S. (1982). *Principle and practice in second language acquisition*. Oxford: Pergamon.
- Lamy, M. y Hampel, R. (2007). *Online communication in language learning and teaching*. England: Palgrave Macmillan.
- Leech, G. (2002). Sobre la importancia de los corpus de referencia. *Donosit*, (24)1, 1-3.
- Lewis, M. (1993). *The lexical approach*. London. England: Language Teaching Publications.
- Lewis, M. (1997). *Implementing the lexical approach*. London, England: Language Teaching Publications.
- Lewis, M. (2000a). *Teaching collocation. Further developments in the lexical Approach*. Hove, England: LTP.

- Lewis, M. (2000b). Materials and resources for teaching collocation. En M. Lewis (Ed.), *Teaching collocation further developments in the lexical approach* (pp.186- 204). Hove, England: LTP.
- Lightbown, P. (1991). What have we here? Some observation on the influence of instruction in L2 learning. En R. Phillipson, E. Kellerman, M. Sharwood-Smith y M. Swam (Eds.), *Foreign/Second language pedagogy research* (pp.197-212). Clevedon, England: Multilingual Matters.
- Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly* (44), 4-30.
- Long, M. H. (1980). *Input, interaction, and second language acquisition* (doctoral dissertation). University of California, Los Ángeles, EE.UU.
- Lozano, M. (2013). Verbos de Apoyo. En F. San Vicente (Ed.), *GREIT. Gramática de referencia de español para italófonos* (pp.801-832). Salamanca, España: Ediciones Universidad de Salamanca.
- Lozano, C. y Mendikoetxea, A. (2013). Learner corpora and second language acquisition: The design and collection of CEDEL2. In Díaz Negrillo, N. Ballier y Thompson (Eds.), *Automatic treatment and analysis of learner corpus data*, Amsterdam: John Benjamins, pp.65-100.
- Lozano, E. y Rodríguez, I. (2014). Las colocaciones léxicas simples del lenguaje jurídico en el derecho civil mexicano. *Debate Terminológico*, 12(1), 31-42.
- Lorenz, G. (1999). Adjective intensification - learners versus native speakers: A corpus study of argumentative writing, Amsterdam: Rodopi.
- Martínez, R. (2013). A framework for the inclusion of multi-word expressions in ELT. *ELT Journal*, 67, 184-198.

Martos, P. (2016). Sobre la aplicación de los criterios de análisis de colocaciones verbo + sustantivo CD a materiales históricos. *ELUA*, 30, 213-238. Doi: 10.14198/ELUA2016.30.10

Mel'čuk, I. (1992). Paraphase et lexique: La Théorie Sens- Texte et le Dictionnaire explicatif et combinatoire. En I. A Mel'čuk,, L. Iordanskaja y S. Mantha (Eds), *Dictionnaire explicatif et combinatoire du français contemporain: Recherches léxico-sémantiques*, Vol IV (pp. 9-58). Montreal, Canada: Les Presses de l'Université de Montréal.

Mel'čuk, I. (1996). Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. En L. Wanner (Ed), *Lexical functions in lexicography and natural language processing* (pp. 37-102). Ámsterdam/ Fildadepfia: John Benjamins.

Mel'čuk, I. (1998). Collocations and lexical functions. En A.P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 23-53). Oxford, England: Clarendon Press.

Mel'čuk, I. y Zolkovskij, A. (1970), Towards a Functioning Meaning-Text Model of Language. *Linguistics* (57), 10-47.

Mendivil, J. (1991). Consideraciones sobre el carácter no discreto de las expresiones idiomáticas. En M. Vide (Ed.). *Actas del VI congreso de lenguajes naturales y formales*. Barcelona, España: PPU.

McCarthy, M. (1990). *Vocabulary*. Oxford, England: Oxford University Press.

Mitchell, T. F. (1971). Linguistic 'going on': collocations and other lexical matters arising on the syntagmatic record. *Archivum Linguisticum*, 2(1), 35- 69.

Monc3, S. (2013). Adquisici3n de las construcciones con el verbo hacer, enfoque plurilingüe. *Revista Nebrija de Lingüística Aplicada*, (13), 1-7.

- Molero, C. y Salazar, D. (2013). Análisis contrastivo, criterios de selección y didácticas de las colocaciones léxicas en el aula de español. En A. Kuzmanović, J. Filipović, J. Stojanović y J. Rajić (Eds.), *Estudios Hispánicos del siglo XXI* (pp.114). Belgrado, Serbia: Universidad de Belgrado.
- Muñiz, E. M^a. (2002). Las colocaciones como combinaciones léxicas libres. En A. V. et al (Eds.). *Léxico y gramática* (pp. 261-271). Lugo, España: Tris Tram.
- Nation, F. (1990). *Teaching and learning vocabulary*. New York, EE.UU.: Newbury House.
- Nation, F. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.
- Nattinger, J. y De Carrico, J.S (1992). *Lexical phrases and language teaching*, Oxford, England: Oxford University Press.
- Navajas, A. (2006). Las colocaciones en el aula de ELE: Actividades para su explotación didáctica. Madrid, España: Universidad Antonio de Nebrija.
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. En J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp- 125-152). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Neuner- Anfindsen, S. (2005). Fremdsprachenlernen und Lernerautonomie. Sprachlernbewusstsein, Lernprozessorganisation und Lernstrategien zum Wortschatzlernen in Deutsch als Fremdsprache. Baltmannsweiler: Schneider Verlag Hohengehren.

- O'Donnell, M. (2008). The UAM Corpus Tool: Software for corpus annotation and exploration. En *Proceedings. Ponencia llevada a cabo en el XXVI Congreso de AESLA, Almería, España*.
- Oportus, R. y Ferreira, A. (2019). Colocaciones gramaticales verbo + preposición en ele: aspectos de naturalidad en aprendientes anglófonos de nivel a2 y b1. *Trabalhos em Linguística Aplicada*, 58(2), 826-858. <https://dx.doi.org/10.1590/010318138653957451621>
- Osorio, G. y Serra, S. (2012). Colocaciones, compuestos sintagmáticos y locuciones nominales: Hacia un intento de delimitación conceptual. *Lenguas modernas*, 39(1), 359-375.
- Parodi, G. (2008). Lingüística de corpus: una introducción al ámbito. *RLA. Revista de lingüística teórica y aplicada*, 46(1), 93-119.
- Pastor, S. (2004). Aprendizaje de segundas lenguas. Lingüística aplicada a la enseñanza de idiomas. Alicante, España: Publicaciones Universidad de Alicante.
- Pawley, A. (2001). Phraseology, linguistics and the dictionary, *International Journal of Lexicography*, 14(2), 122-134.
- Pejovic, A, (2010). Las colocaciones verbales y el contexto. *Colindancias*, (1), 132-140.
- Pérez- Ávila, E. (2006). El corpus lingüístico en la didáctica del léxico en el aula (tesis de máster). Universidad Antonio de Nebrija, Madrid, España.
- Pérez, M. (2014). Análisis de errores colocacionales en un corpus de aprendientes de ELE, MarcoELE. *Revista de Didáctica de Español Como Lengua Extranjera*, (19), 1-20.
- Pérez, M. (2017). Colocaciones con verbos delexicalizados, ¿son más fáciles de aprender? *Language Design*, (19), 127-148.

- Piera, C. y Varela S. (1999). Relaciones entre morfología y sintaxis. En I. Bosque y V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 4367-4421). Madrid, España: Espasa Calpe.
- Pino, A. (2009). Palabras en interacción: un corpus de aprendices suecos de E/LE. A survey of corpus-based research, p. 470-487. Disponible en: <https://goo.gl/XAHPR3>.
- Pino, A. (2012). Saele, un corpus de aprendices suecos de E/LE. In: International conference on corpus linguistics, 4., Jaén. Anais, Murcia: Alienco.
- Porzig, W. (1950). *El mundo maravilloso del lenguaje*. Madrid, España: Gredos.
- Pravec, N. (2002). Survey of learner corpora, *ICAME Journal*, 26, 81-114.
- Prieto, S., Mosqueira, E. y Vásquez, N. (2009). Córpora y enseñanza de lenguas: se buscan colocaciones. En P. Cantos y A. Sánchez (Ed.), *A survey of corpus-based research [Recurso electrónico]* (pp. 366-373). Murcia, España: Asociación Española de Lingüística del Corpus.
- RAE (2009). *Nueva Gramática de la Lengua Española*. Madrid: Espasa Calpe.
- RAE (2018). Diccionario de la lengua española (23.3a ed.). Consultado en <https://www.rae.es>
- Rojo, G. y Palacios, I. (2015). Corpus de aprendices de español (CAES). *Journal of Spanish Language Teaching, Oxford*, 2(2), 194-200, <https://doi.org/10.1080/23247797.2015.1084685>
- Romera, C. (2017). *Construcciones con verbos de apoyo. Caracterización en español y comparativa con el catalán* (tesis de pregrado). Universidad Autónoma de Barcelona, Barcelona, España.

- Ruiz, L. (1997). *Aspectos de fraseología teórica española*. Valencia, España: Servicio de Publicaciones de la Universidad.
- Ruiz, L. (2000). Un enfoque didáctico de la fraseología española para extranjeros. *Quaderns de filologia. Estudis lingüístics*, (5), 259-276.
- Sánchez, A. (2015). La Investigación de Corpus de Aprendientes y el desarrollo de los estudios de la interlengua del español. *Language design: Journal of theoretical and experimental linguistics*, 17, 57-83.
- Shirota, S. (1991). Kotobano en (Solidaridad léxica). Aproximación a la lexicología estructural. Tokio, Editorial Liberta.
- Seco, M. (1978). Problemas formales de la definición lexicográfica. En Universidad de Oviedo (Ed.), *Estudios ofrecidos a E. Alarcos Llorach* (pp.217- 239). Oviedo, España: Universidad de Oviedo.
- Schmitt, N. (2004). *Formulaic sequences*. Ámsterdam, Holland: John Benjamins.
- Shei, C. (2000). An ESL writer's collocational aid. *Computer Assisted Language Learning*, 13(2), 167-182.
- Sidoti (2014). Interferencias colocacionales en construcciones con verbo de apoyo+sustantivo entre lenguas fines. *Revista Lingue e Linguaggi*, (11), 215-224.
- Sinclair, J. (1966). Beginning the study of lexis. En C. Bazell, J. Catford, M. Halliday y R. Robins (Eds.), *In memory of J.R. Firth* (pp. 410-430). Londres, England: Longman.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, England: Oxford University Press.

- Sinclair, J. (2002). The search of units of meaning". En G. Corpas (Ed.), *Las lenguas en Europa: Estudios de Fraseología, Fraseografía y Traducción* (pp.7-37). Granada, España: Comares.
- Soto, R. (2012). Las combinaciones verbo-nominales en las entradas pluriverbales del Diccionario Ejemplificado de Chilenismos (Dech): Una propuesta de categorización desde la perspectiva de la léxico-sintaxis (tesis de maestría). Universidad de Chile, Santiago, Chile.
- Skorepova, A (2008). Estudio tipológico, formal y léxico-semántico de las colocaciones léxicas verbo-nominales en el español actual (tesis de doctorado). Universidad de Granada, Granada, España.
- Thun, H. (1978). Probleme der Phraseologie. Untersuchungen zur wiederholten Rede mit Beispielen aus den Franzosischen, Italienischen, Spanischen und Romiinischen, Beiheftezur Zeitschriftfür Romanische Philologie. Tubinga, Germany: Max Niemeyer.
- Tono, Y. (2003). Learner corpora: Design, development and applications. En D. Archer, P. Rayson, A. Wilson y T. McEnery (Eds.), *Proceedings of the Corpus Linguistics Conference* (pp. 800-809). Lancaster (UK): University Centre for Computer Corpus Research on Language.
- Torjada, M. (2000). Las unidades fraseológicas: Una propuesta didáctica mediante situaciones comunicativas ilustradas en imágenes. *Frecuencia, L*, (15), 27-39.
- Traugott, E. (1978). On the expression of spatio-temporal relations in language", en J. Greenberg, Ch. Ferguson y E. Moravcsik, E. A. (Eds.), *Universals of human language, Vol. III*. Stanford, EE.UU.: Stanford University Press.

- Tristá, A. M^a (1988). *Fraseología y contexto*, La Habana, Cuba: Ed. de Ciencias Sociales.
- Uriel, M. (2014). Las colocaciones en un corpus de aprendices valón y flamenco. En *La enseñanza de ELE centrada en el alumno*. Simposio llevado a cabo en el XXV Congreso Internacional de ASELE, Madrid, España.
- Vinogradov, V. (1947). *Obosnovnij Tipaj fraseologis cheskijedinitis v ruskomyazike*. Moscú.
- Warschauer, M. (2005). Sociocultural perspectives on CALL”, en J. Egbert y G. Mikel Petrie (Eds.) *CALL research and perspectives*. Mahwah, N.J.: Lawrence Earlbaum Associates Inc., pp. 41-51.
- Wen, H. (2019). Introducción al concepto de colocación en la Enseñanza del Español como Lengua Extranjera en China. un acercamiento a español moderno (tomos 1, 2, 3), MarcoELE. Revista de Didáctica de Español Como Lengua Extranjera, (28), 1-13.
- Wotjak, B. (1992). *Verbale Phraseolexeme in System und Text*. Tubinga, Germany: Max Niemeyer Verlag.
- Wotjak, B. (1998). El léxico: encrucijada de la lingüística teórica y aplicada, de metodologías y disciplinas distintas. En J. Fernández y C. Wallhead (Eds.), *Temas de la Lingüística aplicada* (pp. 104-135). Granada, España: Universidad de Granada,
- Zarco, M. (1998). *Predicados complejos y traducción automática*. Cádiz, España: Servicio de Publicaciones.

- Zhang, J. (1993). Extraction of Chinese compound words: an experimental study on a very large corpus. *Annual Meeting of the ACL*, (12), 132-139.
- Zuluaga, A. (1980). Introducción al estudio de las expresiones fijas. (*Studia Romanicaet Linguistica*, 10). Frankfurt-am-Main, Berna, Cirencester: Peter Lang.
- Zuluaga, A. (1990). Unidad fraseológica o fraseologismo. Frankfurt a. M.: Peter de Lang.
- Zuluaga, A. (2002). Los “enlaces frecuentes” de María Moliner: Observaciones sobre las llamadas colocaciones. LEA. *Lingüística Española Actual*, (24)1, 97–114.



13. Anexos

13.1. Etiquetado de colocaciones

Detección de casos de complementos directos

Estimado (a):

Solicito su colaboración para leer un conjunto de textos escritos por estudiantes de Español como Lengua Extranjera ELE con el objetivo de detectar los casos de complementos directos (verbo transitivo+ objeto).

Este material forma parte de un trabajo de investigación cuyos objetivos generales son: 1) Determinar la frecuencia de las colocaciones léxicas simples verbo-nominales de niveles de competencia del ELE A2 y B1 en un sub-corpus del corpus CAELE³² y 2) Comparar las colocaciones léxicas simples verbo-nominales más frecuentes acordes con los niveles de competencia del ELE A2 y B1.

Usted deberá leer cada texto y detectar de forma manual todos los casos en que el estudiante produjo casos de complementos directos, ya sea correctamente o incorrectamente (agramaticales o inexistentes en español). Para ello, 1) subraye el segmento que contiene dicho complemento y 2) señale si considera su uso frecuente o poco frecuente en el español (F/PF).

Su apreciación y comentarios ayudarán a mejorar la rigurosidad, y validez en la identificación de este fenómeno lingüístico, datos necesarios para la investigación señalada.

Antes de comenzar la lectura de los textos, es necesario completar la información solicitada a continuación. Cabe destacar que el uso de los datos aquí recogidos es de carácter confidencial.

³²Corpus de Aprendientes de Español como Lengua Extranjera.

Nombre: _____

Rango etario (marcar con x): **20-25** **25-30**

Carrera: _____

Año

cursado: _____ **Universidad:** _____

Fecha: _____

Gracias por su colaboración.



13.2. Colocaciones léxicas verbo-nominales con verbos de significado pleno en el sub-corpus de ELE

N°	Colocación VP	Frecuencia	%
1	escuchar música	73	15,73%
2	valorar la música	36	7,76%
3	valorar la pintura	28	6,03%
4	valorar la artesanía	22	4,74%
5	valora la literatura	14	3,02%
6	comprar cosas	12	2,59%
7	mantener una tradición	12	2,59%
8	leer un libro	11	2,37%
9	preferir la música	9	1,94%
10	escuchar canciones	9	1,94%
11	escuchar regueton	8	1,72%
12	hacer cosas	8	1,72%
13	tener historia	8	1,72%
14	vender productos	8	1,72%
15	mantener la cultura	7	1,51%
16	tener ganas	7	1,51%
17	tocar música	7	1,51%
18	usar instrumentos	7	1,51%
19	vender cosas	7	1,51%
20	vender artesanía	7	1,51%
21	cantar canciones	6	1,29%
22	comprar productos	6	1,29%
23	difundir leyendas	6	1,29%
24	escribir un libro	6	1,29%
25	valer la pena	6	1,29%
26	ver pinturas	6	1,29%
27	leer un poema	5	1,08%

28	ganar dinero	5	1,08%
29	hacer parte	5	1,08%
30	llamar la atención	5	1,08%
31	tocar un instrumento	5	1,08%
32	hacer música	5	1,08%
33	conocer la cultura	4	0,86%
34	conservar la cultura	4	0,86%
35	contar historias	4	0,86%
36	crear historias	4	0,86%
37	dar cuenta	4	0,86%
38	escribir historias	4	0,86%
39	forman parte	4	0,86%
40	ganar un premio	4	0,86%
41	leer literatura	4	0,86%
42	reflejar la cultura	4	0,86%
43	tratar temas	4	0,86%
44	comprar artesanía	3	0,65%
45	conocer canciones	3	0,65%
46	difundir productos	3	0,65%
47	elaborar la música	3	0,65%
48	escribir un texto	3	0,65%
49	expresar una opinión	3	0,65%
50	hablar un idioma	3	0,65%
51	leer leyendas	3	0,65%
52	mostrar la música	3	0,65%
5	preservar la cultura	3	0,65%
54	producir música	3	0,65%
55	recomendar un lugar	3	0,65%
56	representar una cultura	3	0,65%
57	tener una tradición	3	0,65%

58	tener tiempo	3	0,65%
59	tocar canciones	3	0,65%
60	valora la cultura	3	0,65%
61	valorar los pueblos	3	0,65%
Total	-	464	100%



13.3. Colocaciones léxicas con verbos de significado pleno de nivel A2

Nº	Colocación VP A2	Frecuencia	%
1	escuchar música	35	14,11%
2	valorar la música	21	8,47%
3	valorar la pintura	16	6,45%
4	valorar la artesanía	9	3,63%
5	comprar cosas	8	3,23%
6	mantener una tradición	8	3,23%
7	tener historia	8	3,23%
8	preferir la música	7	2,82%
9	valorar la literatura	7	2,82%
10	usar instrumentos	6	2,42%
11	hacer cosas	5	2,02%
12	tener ganas	5	2,02%
13	tocar música	5	2,02%
14	ver pinturas	5	2,02%
15	cantar canciones	4	1,61%
16	conocer la cultura	4	1,61%
17	crear historias	4	1,61%
18	difundir leyendas	4	1,61%
19	escribir un libro	4	1,61%
20	escuchar regueton	4	1,61%
21	leer literatura	4	1,61%
22	leer un libro	4	1,61%
23	mantener la cultura	4	1,61%
24	valer la pena	4	1,61%
25	contar historias	3	1,21%
26	dar cuenta	3	1,21%
27	escuchar canciones	3	1,21%
28	hablar un idioma	3	1,21%

29	leer un poema	3	1,21%
30	preservar la cultura	3	1,21%
31	tener una tradición	3	1,21%
32	tocar canciones	3	1,21%
33	tocar un instrumento	3	1,21%
34	valora la cultura	3	1,21%
35	comprar artesanía	2	0,81%
36	escribir un texto	2	0,81%
37	hacer música	2	0,81%
38	hacer parte	2	0,81%
39	llamar la atención	2	0,81%
40	recomendar un lugar	2	0,81%
41	reflejar la cultura	2	0,81%
42	tener tiempo	2	0,81%
43	tratar temas	2	0,81%
44	vender cosas	2	0,81%
45	vender productos	2	0,81%
46	conocer canciones	1	0,40%
47	conservar la cultura	1	0,40%
48	escribir historias	1	0,40%
49	expresar una opinión	1	0,40%
50	forman parte	1	0,40%
51	ganar dinero	1	0,40%
52	ganar un premio	1	0,40%
5	mostrar la música	1	0,40%
54	producir música	1	0,40%
55	representar una cultura	1	0,40%
56	valorar los pueblos	1	0,40%
57	comprar productos	0	0,00%
58	difundir productos	0	0,00%

59	elaborar la música	0	0,00%
60	leer leyendas	0	0,00%
61	vender artesanía	0	0,00%
Total	-	248	100,00%



13.4. Colocaciones léxicas con verbos de significado pleno de nivel B1

Nº	Colocación VP B1	Frecuencia	%
1	escuchar música	38	17,59%
2	valorar la música	15	6,94%
3	valorar la artesanía	13	6,02%
4	valorar la pintura	12	5,56%
5	leer un libro	7	3,24%
6	valora la literatura	7	3,24%
7	vender artesanía	7	3,24%
8	comprar productos	6	2,78%
9	escuchar canciones	6	2,78%
10	vender productos	6	2,78%
11	vender cosas	5	2,31%
12	comprar cosas	4	1,85%
13	escuchar regueton	4	1,85%
14	ganar dinero	4	1,85%
15	mantener una tradición	4	1,85%
16	conservar la cultura	3	1,39%
17	difundir productos	3	1,39%
18	elaborar la música	3	1,39%
19	escribir historias	3	1,39%
20	forman parte	3	1,39%
21	ganar un premio	3	1,39%
22	hacer cosas	3	1,39%
23	hacer música	3	1,39%
24	hacer parte	3	1,39%
25	leer leyendas	3	1,39%
26	llamar la atención	3	1,39%
27	mantener la cultura	3	1,39%
28	cantar canciones	2	0,93%

29	conocer canciones	2	0,93%
30	difundir leyendas	2	0,93%
31	escribir un libro	2	0,93%
32	expresar una opinión	2	0,93%
33	leer un poema	2	0,93%
34	mostrar la música	2	0,93%
35	preferir la música	2	0,93%
36	producir música	2	0,93%
37	reflejar la cultura	2	0,93%
38	representar una cultura	2	0,93%
39	tener ganas	2	0,93%
40	tocar música	2	0,93%
41	tocar un instrumento	2	0,93%
42	tratar temas	2	0,93%
43	valer la pena	2	0,93%
44	valorar los pueblos	2	0,93%
45	comprar artesanía	1	0,46%
46	contar historias	1	0,46%
47	dar cuenta	1	0,46%
48	escribir un texto	1	0,46%
49	recomendar un lugar	1	0,46%
50	tener tiempo	1	0,46%
51	usar instrumentos	1	0,46%
52	ver pinturas	1	0,46%
5	conocer la cultura	0	0,00%
54	crear historias	0	0,00%
55	hablar un idioma	0	0,00%
56	leer literatura	0	0,00%
57	preservar la cultura	0	0,00%
58	tener historia	0	0,00%

59	tener una tradición	0	0,00%
60	tocar canciones	0	0,00%
61	valorar la cultura	0	0,00%
-	Total	216	100,00%

