

COLOCACIONES GRAMATICALES VERBO + PREPOSICIÓN EN ELE: ASPECTOS DE NATURALIDAD EN APRENDIENTES ANGLÓFONOS DE NIVEL A2 Y B1¹

GRAMMATICAL COLLOCATIONS VERB + PREPOSITION IN SPANISH AS A FOREIGN LANGUAGE: ASPECTS OF NATURALNESS IN A2 AND B1 ANGLOPHONE LEARNERS

René Edgardo Oportus Torres*
Anita Alejandra Ferreira Cabrera**

RESUMEN

El siguiente trabajo presenta los resultados del análisis de dos sub-corpus escritos de aprendientes anglófonos de ELE de niveles A2 y B1, cuyo propósito es examinar la producción de colocaciones gramaticales verbo + preposición (ColGram v+p) en relación con la producción de un corpus comparable de hablantes nativos (HN) de español. Para ello, se identifican las ocurrencias colocacionales, se determina su corrección desde un punto de vista de la combinatoria colocacional, y se describen los usos aceptables acorde el patrón sintáctico-semántico que las caracteriza. 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. Lo anterior se explicaría en cuanto se constata el uso recurrente de algunas colocaciones en particular en ambos grupos de aprendientes, las que contribuyen a engrosar la frecuencia; no así, la variedad. Además, en los usos correctos sobresalen las frecuencias inferiores de los aprendientes en los subtipos de ColGram v+p que incluyen usos pronominales de verbos, contrastando con el mayor uso de los HN. Este trabajo aporta desde un punto fraseológico a la investigación de las combinaciones verbo + preposición de fijación intermedia de la interlengua colectiva de aprendientes anglófonos de ELE.

Palabras-clave: colocaciones gramaticales verbo + preposición, análisis contrastivo de la interlengua, corpus de aprendientes electrónico.

*Universidad de Concepción, Chile. reneoportus@udec.cl.

** Universidad de Concepción, Chile. afferreir@udec.cl, <https://orcid.org/0000-0001-7979-6467>

1. Este estudio forma parte de la tesis doctoral "Influencia de la L1 en el Procesamiento y Errores en el Uso de Colocaciones Gramaticales en ELE" de René Oportus Torres bajo la dirección de la Dra. Anita Ferreira Cabrera. Este artículo se ha desarrollado en el contexto del proyecto de investigación FONDECYT N° 1180974 "Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua", cuya investigadora responsable es la Dra. Anita Ferreira Cabrera.

ABSTRACT

The following article presents the results of the analysis of two written sub-corpora of Anglophone learners of Spanish a foreign language of levels A2 and B1, whose objective is to examine the production of grammatical collocations verb + preposition (GramCol v+p) in relation with the production of a comparable corpus of native speakers (NS) of Spanish. To this end, collocational occurrences are identified, correction from a collocational combinatory viewpoint is determined, and acceptable uses are described according to characterizing syntactic-semantic patterns. Based on frequency data, it is concluded that learners exhibit features of restricted naturalness in the use of the observed units, in terms that A2 level produces a lower number and lower variety than NS; while B1 learners, though reaching similar frequency as NS, they also present a lower variety, similar to that of A2. The above could be explained with the recurrent use of some collocations in particular by both groups of learners which contribute to gross frequency, though not to variety. Additionally, within acceptable collocational uses, learners' lower frequency of subtypes which include pronominal verbs stands out as a difference regarding NS use. This work contributes from a phraseological perspective to the study of combinations verb + preposition of intermediate fixedness in the collective interlanguage of Anglophone learners of Spanish as a foreign language.

Keywords: grammatical collocations verb + preposition, interlanguage contrastive analysis, learners' corpus in electronic format.

INTRODUCCIÓN

El desarrollo de la competencia fraseológica, referida al uso de combinaciones de palabras en la forma de expresiones semi-prefabricadas de la lengua, ha generado un gran interés en la investigación en ASL (GRANGER Y BESTEN, 2014). Entre las expresiones estudiadas se distinguen varios subtipos definidos de acuerdo a criterios de composicionalidad, significado metafórico y fijación sintáctica (colocaciones, locuciones idiomáticas, paremias, entre otros). Tradicionalmente, la investigación se ha focalizado en las unidades que muestran un alto nivel de fijación, en desmedro de aquellas caracterizadas por un mayor grado de composicionalidad y sustitución, y por rasgos de fijación intermedia entre sus elementos, como es el caso de las colocaciones (HENRIKSEN, 2013). No obstante, los resultados de la investigación han apuntado a una mayor frecuencia y recurrencia de estas últimas, lo que destaca su relevancia (GRANGER, PAQUOT Y RAYSON, 2008; RICA, 2010; HENRIKSEN, 2013; GYLLSTAD Y WOLTER, 2015).

Un área particular de la temática colocacional que ha recibido escasa atención en la investigación de ELE es el estudio del grupo de las colocaciones gramaticales (por ejemplo, las combinaciones verbo + preposición), comparado con otras expresiones fraseológicas, o con el grupo particular de las colocaciones léxicas (TRAVALIA, 2006; DOMÍNGUEZ, 2014; FERNÁNDEZ, 2014). Esto derivaría,

en primer lugar, de los múltiples planteamientos terminológico-conceptuales disponibles en la literatura especializada en torno a qué características definen a las colocaciones en general, lo que ha conducido a cierta confusión terminológica (SÁNCHEZ, 2010; GYSLTAD Y WOLTER, 2015). En segundo, lugar aparece el cuestionamiento sobre si el grupo específico de colocaciones gramaticales debe ser incluido y estudiado desde la perspectiva fraseológica, o bien, si estas corresponderían a fenómenos ya abordados por el estudio de otras parcelas lingüísticas; por ejemplo, el caso de régimen preposicional (FERNÁNDEZ, 2014; VINCZE, 2015).

No obstante lo anterior, en ASL existe un amplio acuerdo sobre la importancia del desarrollo de la competencia colocacional en los aprendientes, respecto del manejo y comprensión de las colocaciones de la L2, puesto que esta contribuye en forma importante a los aspectos de naturalidad y fluidez en la comunicación, y a la eficiencia en el procesamiento de la L2 (CHEIKH-KHAMIS, 2013). Así también, en un ámbito más orientado hacia la didáctica, se recalca que conocer una palabra implica no solo saber sobre sus diferentes acepciones, usos, y áreas temáticas en las que ocurre; sino también los contextos lingüísticos en que aparece; es decir, los términos con los que combina frecuentemente (LU, 2016).

El presente trabajo se enmarca dentro de una investigación mayor cuyo objetivo principal es examinar el procesamiento cognitivo de las colocaciones gramaticales verbo + preposición (ColGram v + p) en aprendientes anglófonos de ELE desarrollada en el Programa ELE-UdeC², y su motivación se origina en los hallazgos obtenidos a partir de las investigaciones internas que preceden a este trabajo. Específicamente, se ha evidenciado que uno de los problemas de mayor frecuencia y recurrencia en los aprendientes anglófonos es el uso preposicional (FERREIRA, ELEJALDE Y VINE, 2014; FERREIRA Y LAFLEUR, 2015; FERREIRA, OPORTUS Y FUENTES, 2016; FERREIRA Y ELEJALDE, 2017; FERREIRA Y OPORTUS, 2018). En esta senda, nos proponemos investigar dicho problema desde una mirada específica y selectiva, adoptando una perspectiva fraseológica para examinar el uso de concurrencias verbo + preposición, en donde el sintagma preposicional exigido por el verbo para completar su significado no constituye un caso de régimen obligatorio. Siguiendo a Travalia (2006), estas secuencias corresponderían a ColGram v+p, cuya definición y tipología adhieren al concepto básico de colocación referido a la ocurrencia frecuente, pero no obligatoria, de un término léxico y una preposición; las que a su vez también difieren de las combinaciones libres (BUYSE, 2003).

2. Programa de Español como Lengua Extranjera. Universidad de Concepción (Ele-udec.cl)

La metodología del estudio presentado se basa en el modelo de Análisis Contrastivo de la Interlengua (ACI) y los procedimientos de la investigación de Corpus Electrónicos de Aprendientes para examinar la frecuencia de uso de ColGram v+p en la producción escrita de estudiantes anglófonos de ELE de nivel A2 y B1 en comparación con la producción de un corpus comparable de HN de español. Nos interesa indagar en el aspecto de naturalidad y expresividad de los aprendientes en relación con los HN, en términos de los rasgos similares y distintos en la producción de ColGram v+p. Para ello, conceptualizamos a estas unidades como combinaciones de fijación intermedia, e identificamos los usos correctos e incorrectos, clasificándolos de acuerdo a una tipología basada en un criterio sintáctico-semántico (TRAVALLIA, 2006). En definitiva, intentamos responder a las interrogantes de cómo difieren los aprendientes y HN respecto de (1) la frecuencia y variación del uso de ColGram v+p, (2) la cantidad de usos correctos e incorrectos producidos, y (3) la frecuencia de usos correctos acorde el patrón y subtipos sintáctico-semánticos.

El artículo se estructura de la siguiente manera. En primer lugar, el marco teórico introduce: las diversas propuestas terminológicas y conceptuales en torno al estudio colocacional, la perspectiva de análisis y propuestas tipológicas consideradas para el estudio de las ColGram v+p, así como también los antecedentes disponibles en la literatura especializada. En segundo lugar, se describe el estudio con su marco metodológico, datos fuente y muestra examinada. Posteriormente, se incluye la sección de análisis, los resultados y su discusión. Finalmente, se presentan las principales conclusiones, y limitaciones del estudio.

1. MARCO TEÓRICO

1.1. El concepto de colocación

El concepto de colocación es un término de naturaleza difusa y compleja. De ello dan cuenta las numerosas publicaciones que han intentado clarificar los aspectos terminológicos y operacionales en base al contraste con otras unidades lingüísticas (HIGUERAS, 2004; SÁNCHEZ, 2010). Este debate también ha incluido numerosas propuestas en torno a definir qué marco disciplinario debería encargarse primariamente de su estudio. En este sentido, como la concepción de colocación implica relaciones en el plano sintagmático, la demarcación del alcance de los campos de la sintaxis y la fraseología se torna difusa (GRANGER Y PAQUOT, 2008); por ejemplo, BOSQUE (2001) sostiene que estas unidades

deberían situarse bajo el foco intermedio de la interfaz léxico-sintaxis. Obviamente, los diferentes matices y perspectivas respecto de la conceptualización adoptada tendrían una implicancia directa en los resultados de la investigación colocacional.

No obstante lo anterior, dos enfoques han prevalecido en la investigación colocacional: el enfoque estadístico, y el fraseológico. En el primer caso, el criterio principal para identificar las colocaciones se basa en la frecuencia reiterada con la que ocurren las combinaciones de palabras, lo que se estima aplicando medidas de asociación (GYLLSTAD Y WOLTER, 2015). Las mayores críticas a este enfoque se refieren a que la co-aparición frecuente de dos términos no determina por sí sola la ocurrencia de una colocación, y a que dichos datos no aportan información sobre la naturaleza lingüística de las combinaciones de interés (SÁNCHEZ, 2010; FERNÁNDEZ, 2014). Por otra parte, el enfoque fraseológico utiliza como criterio fundamental en la identificación de las unidades colocacionales la tipología respecto de los aspectos de estructura gramatical y grado de transparencia semántica (GYLLSTAD Y WOLTER, 2015). Sobresale en esta perspectiva la distinción del estatus de los constituyentes colocacionales entre base y colocativo, en donde la base determina semánticamente la selección del colocativo.

Adicionalmente a estos dos paradigmas, desde un punto de vista esencialmente semántico se propone que las unidades colocacionales corresponden a casos de restricciones sistemáticas intralingüísticas durante el proceso de selección léxica, las que se establecen acorde con la propiedades del predicado, sea este un sustantivo, verbo u otro elemento léxico (SÁNCHEZ, 2010). En este sentido, no bastaría con el conocimiento del mundo, y del significado de un ítem léxico para ser capaz de deducir los elementos lingüísticos correspondientes a su extensión. Sin duda, esta visión subraya la relevancia del desarrollo de la competencia colocacional en el contexto de ASL.

Actualmente en ASL es frecuente incluir los estudios colocacionales bajo el prisma de la fraseología y señalar que estas unidades se sitúan en un punto intermedio entre las locuciones y las combinaciones libres. (FERNÁNDEZ, 2014). Al respecto, Buyse (2003) contrasta las propiedades de las colocaciones versus las locuciones, indicando que las primeras (1) establecen una relación típica entre base y colocativo, (2) muestran menor grado de fijación que las locuciones, incluyendo posibles transformaciones, sustituciones y adiciones, (3) expresan un significado relativamente literal y (4) transparente. Por otra parte, las colocaciones diferirían de las combinaciones libres en cuanto a que (1) en estas últimas la co-aparición de los constituyentes es resultado de la casualidad, (2) no expresan una relación

típica, (3) conllevan un significado nítido, (4) presentan menores restricciones de combinatoria, y (4) no son identificadas por los hablantes como una unidad.

1.2. Las colocaciones gramaticales en ELE

En ELE la identificación de la categoría de colocaciones gramaticales ha sido fuertemente influenciada por la distinción de Benson, Benson e Ilson (1986) respecto de la lengua inglesa. Los autores clasifican a las unidades colocacionales en *colocaciones léxicas*, dependiendo de si la combinación incluye una secuencia de dos elementos léxicos, o en *colocaciones gramaticales*, si estas comprenden una secuencia de una palabra léxica y una gramatical, por ejemplo, *verbo + preposición*, o *adjetivo + pronombre relativo*. Sin embargo, aunque esta postura es acogida por varios lingüistas; también ha sido criticada en base a que las colocaciones gramaticales atañerían al campo de la sintaxis; por ejemplo, situándolas dentro del ámbito de los sellos gramaticales de las palabras (restricciones sintácticas en el uso del léxico), o de los patrones de valencias (GRANGER Y PAQUOT, 2008). Así también, VINCZE (2015) alude a que las colocaciones gramaticales serían abordadas por el estudio de otros fenómenos lingüísticos, como es el caso de los patrones de régimen de unidades léxicas específicas.

No obstante los cuestionamientos anteriores, las tipologías colocacionales para el español han tendido a incluir a las ColGram dentro de sus categorías. Este es el caso de Corpas (1996), quien contempla en el tercer grupo de su tipología colocacional al tipo sustantivo + preposición + sustantivo; y a Koike (2001), quien incorpora en el grupo 3 de su propuesta a las combinaciones verbo + preposición + sustantivo (por ejemplo, jugar al fútbol, pagar con tarjeta, ir en coche, trabajar / hacer de canguro, etc.). Lo mismo, ha ocurrido con trabajos posteriores en el ámbito de ELE (BUYSE, 2003; GONZÁLEZ, 2006; DOMÍNGUEZ, 2014; LU, 2016). En este contexto, los planteamientos de Carolina Travalía (2006) resultan muy significativos para la realización de este trabajo, en cuanto apuntan a precisar la conceptualización de las ColGram para el español y especificar diferentes subtipos y consideraciones dentro de dicha categoría.

La investigadora mencionada aboga por un replanteamiento sobre lo que se ha entendido por ColGram, insistiendo en el criterio esencial de fijación intermedia para su identificación; en contraste con la ocurrencia obligatoria de un término léxico y uno gramatical, casos que tradicionalmente han sido clasificados como ColGram. De esta forma, descarta considerar secuencias tales como *depende de* en esta categoría debido a su fijación eminentemente obligatoria en la lengua. Así también, subraya la importancia del contraste entre las ColGram y combinaciones

libres. En este sentido, el contexto facilitaría la identificación de las ColGram en construcciones tales como *escribir sobre una máquina*; en donde la interpretación de ColGram estaría dada por el significado de *tema sobre el que se escribe* versus una combinación libre correspondiente al significado de *lugar sobre el que se realiza la acción* (identificado como CC³ en RAE, 2009). En definitiva, Travalía (2006) propone 6 tipos de colocaciones gramaticales, dentro de los cuales, dos de ellos corresponden al tipo ColGram v+p.

1.3. Corpus de Electrónicos de aprendientes, ACI y AE

La fuente de datos principal para realizar la investigación en ASL está constituida por la lengua misma producida por los aprendientes; ya sea, obtenida mediante procedimientos experimentales, juicios metalingüísticos (por ejemplo, protocolos de pensar en voz alta), o mediante su producción en condiciones naturales (GRANGER, 2002). La investigación de Corpus de Aprendientes y el desarrollo de los corpus de aprendientes electrónicos se han constituido en procedimientos claves para el acceso a datos naturales de la interlengua, los que permiten describir diferentes estadios en su desarrollo y determinar los factores que afectan a este proceso (SÁNCHEZ, 2015a). A partir de estas herramientas, se desarrolla el modelo investigativo denominado Análisis Contrastivo de la Interlengua (ACI) que trata a la lengua de los aprendientes como una entidad autónoma e independiente (GRANGER, 2015), y permite mediante su enfoque cuantitativo y cualitativo realizar comparaciones sistemáticas entre grupos de aprendientes y hablantes nativos.

El ACI comprende la realización de diferentes tipos de contrastes para caracterizar la interlengua colectiva de los aprendientes de la manera más completa posible. Estos incluyen la comparación entre variedades de la interlengua, por ejemplo, entre aprendices de distintas lenguas maternas, o entre grupos con diferentes niveles de competencia, además de la incorporación de otros factores inherentes al proceso de adquisición (motivación, edad, etc.). Así también, se concibe esencial la comparación entre aprendientes y HN tomando como condición que las fuentes de datos sean comparables. Este último tipo de contraste, correspondiente al empleado en este estudio, se realiza con el objeto de determinar el grado de naturalidad y expresividad respecto de la variante experta, identificando semejanzas, o diferencias, tales como los usos infrarrepresentados o sobrerrepresentados (SÁNCHEZ, 2015a). Por último, a estas comparaciones, también se integra el análisis de los

3. Complemento circunstancial

errores de aprendientes asistido por el computador, de tal forma de dar cuenta en forma amplia de los aciertos y dificultades de los aprendientes (FERREIRA, 2018).

1.4. Las ColGram y la investigación en ELE

La investigación sobre el manejo y adquisición de ColGram se ha desarrollado principalmente en el inglés como L2; lo que contrasta con el ámbito de ELE, en donde contamos con datos escasos sobre estas unidades. A saber, Domínguez (2014) es el único trabajo disponible en ELE, el que compara la producción general de colocaciones léxicas y gramaticales entre aprendientes de L1 valón⁴ y de L1 neerlandés. Los resultados señalan que el número de ColGram supera a las colocaciones léxicas en ambos grupos, siendo el tipo léxico en donde se produce la mayor variedad de unidades. Por otra parte, en los estudios en L2 inglés, Rica (2010) reporta que los aprendientes de nivel B1 y B2 producen más colocaciones gramaticales (varios tipos, incluyendo ColGram v+p) que los HN. Mientras que Bahardoust y Moeni (2012) informan que la proporción de ColGram v+p producidas correctamente supera a las incorrectas en estudiantes de inglés de nivel indeterminado (posiblemente intermedio).

En definitiva, como señalan algunos autores (TRAVALIA, 2006; RICA, 2010; DOMÍNGUEZ, 2014), de acuerdo a la revisión de la literatura especializada, pudimos constatar que en ELE existen datos escasos sobre la temática de las ColGram en general; por ejemplo, en lo que se refiere al manejo de ColGram por parte de los aprendientes versus los HN. A partir de esta observación, Rica (2010) apunta a la necesidad de emprender estudios específicos relativos al grupo de ColGram v+p, pues estas constituyen el grupo más numeroso dentro de las ColGram. Por otra parte, Henriksen (2013) aboga por la realización de investigaciones de mayor especificidad en torno a la necesidad de describir en forma sistemática las subcategorías que pueden distinguirse dentro de los grupos de ColGram más amplios ya identificados.

2. EL ESTUDIO

El diseño de esta investigación es descriptivo y transversal, con un análisis mixto de los datos. Vale decir, estos se procesan cuantitativamente en base a la frecuencia, y cualitativamente durante la identificación, categorización, e interpretación del grado de corrección de las ColGram v+p; proceso marcado

4. El valón es una lengua de la misma rama que el francés y se habla en el sur de Bélgica.

por la naturaleza idiosincrática de estas unidades (RODRÍGUEZ-FERNÁNDEZ, CARLINI Y WANNER, 2015). El objetivo principal es comparar la frecuencia de usos correctos e incorrectos, variedad de unidades producidas, y frecuencia a lo largo de patrones sintáctico-semánticos entre los aprendientes de ELE y la producción de los HN. En definitiva, se intenta responder a las preguntas señaladas a continuación.

2.1. Preguntas de investigación

1. ¿Cómo difieren la frecuencia y variedad de ColGram v+p en ambos niveles de competencia de estudiantes extranjeros con respecto a la producción de los HN?
2. ¿Cómo difieren los aprendientes extranjeros de los HN en la frecuencia de usos correctos de ColGram v+p según los patrones sintáctico-semánticos que describen a estas unidades?

2.2. El Corpus de aprendientes y de HN

Tanto el sub-corpus de aprendientes y de HN sobre los que se basa el estudio han sido desarrollados e implementados en el marco de los proyectos de investigación FONDECYT N° 1140651 (FERREIRA, 2014) y N° 1180974 *"Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua"* (FERREIRA, 2018). El Corpus especializado de Aprendientes de Español como Lengua Extranjera CAELE cuenta con 729 textos (hasta el año 2017) redactados por aprendientes de diversos orígenes lingüísticos (FERREIRA Y ELEJALDE, 2017), mientras que el Corpus de HN de español contiene 164 textos (FERREIRA, 2014-2018).

En consideración de los objetivos del estudio, en el caso de los aprendientes se seleccionaron solo los textos de los 33 estudiantes de lengua materna inglés para conformar el sub-corpus de aprendientes anglófonos. Su nivel de competencia (A2=19; B1=14) fue determinado con la Prueba de Multinivel con Fines Específicos Académicos del Programa ELE-UdeC (FERREIRA, 2016), instrumento alineado con los niveles del Marco Común Europeo de Referencia (CONSEJO DE EUROPA, 2001), aplicado al comienzo del programa. Las actividades de escritura se realizaron en las clases de español, directamente en el computador bajo la supervisión de un docente, sin diccionarios ni correctores. En total, el tamaño del sub-corpus ascendió a 207 textos (116 textos en nivel A2, y 91 en B1), con 24790 palabras en el nivel A2 y 24789 en B1 (extensión promedio de 193, y 244 palabras

por texto, respectivamente). La elicitación de los textos estuvo dirigida por una función comunicativa, y estos fueron redactados en estilo narrativo, descriptivo y argumentativo. Los temas incluyeron aspectos culturales y problemáticas de la sociedad chilena y latinoamericana, con un foco en la apreciación de la realidad de los países de los estudiantes. La cantidad total de temas cubiertos fue de 20 en A2, y 16 en B1 en las temáticas dadas en la tabla 1.

Tabla 1. Número de textos por temática y por nivel

Temática	N° textos por temática		
	A2	B1	Total
1. Artesanías chilenas	9	10	19
2. Artesanías del país de origen	3	4	7
3. Dieta vegetariana y carnívora	5	-	5
4. Experiencias en las vacaciones	5	-	5
5. La familia como núcleo social	5	-	5
6. Literatura chilena	1	5	6
7. Literatura del país de origen	8	9	17
8. Leyendas chilenas y micro relatos	4	4	8
9. Lugares turísticos chilenos	5	6	11
10. Lugares turísticos latinoamericanos	3	6	9
11. Lugares turísticos del país de origen	4	6	10
12. Música chilena	9	10	19
13. Música latinoamericana	4	4	8
14. Música del país de origen	8	8	16
15. Música favorita y música juvenil	8	3	11
16. Ocio y diversión en el país de origen	6	-	6
17. Pintura chilena	7	5	12
18. Problemas juveniles	8	4	12
19. Problemas socio-ambientales	7	4	11
20. Pueblos originarios	7	3	10
N° total textos	116	91	207

Para poder contar con un corpus control apropiado para el estudio, correspondiente al Sub-corpus de HN, se seleccionaron a los participantes que contaban con la realización de todas las tareas de escritura en el corpus general de HN, cuya lengua materna era el español. En definitiva, el corpus control se

implementó con 72 textos correspondientes a tres tareas de escritura redactadas por 24 estudiantes de primer año de periodismo de una universidad chilena, con un total de 24313 palabras (extensión promedio = 303 palabras por texto). Las tareas se realizaron directamente en el aula virtual de la asignatura, y todas ellas correspondieron a una primera instancia de escritura, antes que los estudiantes recibieran algún tipo de feedback correctivo. Cada estudiante redactó tareas en tres sub-géneros periodísticos de *opinión* (PÁEZ, 2004): (1) una en el modelo de *columna de opinión*, de temática libre (2) una en el modelo de *comentario*, de temática libre, y (3) una en el modelo de *opinión*, de tema común para todos los estudiantes. En cuanto a los estilos discursivos, estos corresponden a la modalidad expositiva y argumentativa.

2.3. Metodología de la investigación

El estudio se sustenta en el enfoque de investigación del ACI (GRANGER, 2015) y en los procedimientos metodológicos del Análisis de Corpus Electrónicos de Aprendientes desarrollados en los proyectos FONDECYT N° 1140651 (FERREIRA, 2014) y N° 1180974 (FERREIRA, 2018). La identificación, clasificación, y anotación de las instancias de ColGram v+p correctas o incorrectas se efectuó mediante la utilización de la herramienta UAM Corpus Tool versión 3.3 (O'DONNELL, 2008).

2.4. Conceptualización y tipología de ColGram v+p del estudio

En este trabajo las ColGram están conceptualizadas sobre la base de las características asociadas tradicionalmente a las colocaciones léxicas (TRAVALIA, 2006). En términos operacionales las definimos como la *combinación de un verbo pleno y una preposición (o locución preposicional) que introduce un complemento preposicional no obligatorio*. En otras palabras, estos complementos cumplirían la función sintáctica de un grupo argumental seleccionado semánticamente por el verbo, pero sintácticamente puede omitirse al estar implícitos, o pueden alternar con otros grupos preposicionales o sintácticos para expresar un sentido similar.

Siguiendo los planteamientos teóricos respecto de la ColGram y luego de la revisión preliminar de los sub-corpus, se desarrolló una tipología específica de análisis de las ColGram v+p. Basados en la propuesta de Travalía (2006) y los rasgos de la función sintáctica del complemento preposicional, correlato gramatical de nuestras unidades de estudio (RAE, 2009; JIMÉNEZ, 2011; VAISBERG, 2013), se identificaron cuatro patrones sintáctico-semánticos principales, con subtipos

específicos dentro de ellos, de acuerdo con (1) la alternancia de preposiciones distintas después del verbo para expresar un significado similar, (2) la rección de la preposición sobre el término preposicional respecto de la exigencia de una estructura sintáctica específica a continuación de la preposición, (3) la alternancia del complemento con otras funciones sintácticas, y (4) la posibilidad de supresión del complemento preposicional en algunos contextos.

A continuación, se presentan las 4 categorías de patrones sintáctico-semánticos generales identificados en la tipología de ColGram v+p del estudio y los subtipos dentro de ellos. Estos últimos se ilustran con ejemplos tomados de los sub-corpus analizados:

- **Patrón 1.** Verbo + 2 ó más preposiciones alternantes con un significado similar.
 - Subtipo 1.1 Verbo + (CD⁵) + CPrep⁶ con preposiciones alternantes: **Hay mucha gente que está hablando sobre Manual Garcia.*⁷
 - Subtipo 1.2 Verbo + CPrep con preposiciones alternantes: *Fuimos a un festival de cerveza artesanal.*
 - Subtipo 1.3 Verbo pronominal + CPrep con preposiciones alternantes: *Se trata sobre un viaje a Las Vegas.*
- **Patrón 2.** Verbo + 1 ó más de 1 preposición + estructura sintáctica específica.
 - Subtipo 2.1 Verbo + CPrep con estructura sintáctica específica (ej., infinitivo): **Y aprendes a aplicar esos ideas.*
 - Subtipo 2.2 Verbo + CD + CPrep con estructura sintáctica específica: *Para apoyar a los músicos chilenos a difundir mejor su música.*
 - Subtipo 2.3 Verbo + sintagma preposicional predicativo del sujeto o del CD: **Me llevaría ropa y cerveza a recuerdo (Me llevaría ropa y cerveza de recuerdo).*

5. Complemento directo

6. Complemento preposicional

7. Los ejemplos provienen de los datos obtenidos del análisis de los sub-corpus de estudiantes anglófonos CAELE y del sub-corpus de HN. Los segmentos subrayados señalan la ocurrencia de la unidad colocacional. El asterisco (*) que precede a los enunciados indican la presencia de algún error en diversos niveles de análisis lingüístico desde un punto de vista prescriptivo; por ejemplo, de concordancia, ortográficos u otros,

• **Patrón 3.** Verbo + CPrep alternante con CD con diferencias semánticas leves o marcadas entre las alternancias.

Subtipo 3.1 Verbo no pronominal + CPrep que alterna con verbo no pronominal + CD con significados cercanos: *Indica como uno de sus objetivos el acabar con la violencia (acabar algo versus acabar con algo).*

Subtipo 3.2 Verbo no pronominal + CPrep que alterna con verbo no pronominal + CD, con diferencias semánticas importantes: **Creían en la Puchamama. la madre de la tierra (creer algo versus creer en algo).*

Subtipo 3.3 Verbo pronominal + CPrep que alterna con verbo no pronominal + CD, con significados cercanos: *Se encuentra con un desconocido (encontrarse con alguien versus encontrar a alguien).*

Subtipo 3.4 Verbo pronominal + CPrep que alterna con verbo no pronominal + CD, con diferencias semánticas importantes: **Cada artesano se enfoca en algo (enfocarse en algo versus enfocar algo).*

Patrón 4. Verbo + CPrep que puede aparecer tácito.

Subtipo 4.1 Verbo no pronominal + CD + CPrep que puede aparecer tácito: *Obras lindas literarias que deberíamos compartir con el mundo (versus obras lindas literarias que deberíamos compartir).*

Subtipo 4.2 Verbo no pronominal + CPrep que puede aparecer tácito: **Mi preferencia de musica cambia con las emociones mas fuerte en mi vida (versus mi preferencia de música cambia en mi vida).*

Subtipo 4.3 Verbo pronominal + CPrep que puede aparecer tácito: **Chilenos les gusta divertirse mucho con carrete y aventura (versus a los chilenos les gusta divertirse mucho).*

2.5. Identificación y determinación de la corrección de las ColGram v+p

Siguiendo los aspectos de combinatoria, de exigencia semántica y sintáctica, incluidos en la definición operacional del estudio, la identificación de las unidades colocacionales se realizó manualmente en los sub-corpus analizados. No obstante, debido a la forma en que estas secuencias se solapan con otras secuencias de v+p no colocacionales de la lengua, señalamos formalmente las concurrencias v+p que no constituyen instancias de ColGram v+p, por lo cual son excluidas del análisis de este estudio. Estas son las secuencias v+p que se presentan dentro de la estructura interna de locuciones verbales y cuasi-locuciones verbales con

verbos de apoyo, los marcadores funcionales del CD, CI⁸ y complemento agente, CPrep de régimen obligatorio, perífrasis verbales, complementos argumentales de ubicación, complementos preposicionales adjuntos, oraciones copulativas, y las secuencias verbo + locuciones adverbiales introducidas por preposición + CD + complemento del nombre.

La clasificación de las unidades colocaciones según su corrección en correctas o incorrectas se basó en el criterio de combinatoria esperada entre verbo y preposición. Se consideraron correctos aquellos casos en que la preposición producida correspondía a la esperada (o una de las esperadas) según el uso en la lengua para expresar el sentido dado por el contexto del enunciado producido. Este proceso se basó en dos fuentes: (1) la introspección de los investigadores, hablantes nativos de español, respecto del análisis y la interpretación de los datos en base a su conocimiento interiorizado de la lengua y fenómeno de estudio; (2) los datos en línea del Corpus del Español del Siglo XXI CORPES⁹ (RAE, 2018), corpus de referencia representativo de la norma panhispánica del español actual (SÁNCHEZ, 2015b).

Asimismo, debido a la naturaleza fraseológica del estudio, adoptamos el enfoque de Howarth (1996, en VINCZE, 2015) en el caso de las ColGram v+p incorrectas, focalizándonos exclusivamente en los errores que rompían el vínculo colocacional; ya fuese, debido a la omisión o reemplazo de la preposición, o debido a la creación de una combinación v+p de naturaleza colocacional incorrecta en la lengua mediante la adición o falsa selección de una preposición. En consecuencia, excluimos del análisis los errores relativos a otros aspectos lingüísticos que aparecen manifiestos en el contexto de los enunciados donde ocurría la ColGram v+p, por ejemplo, a nivel de la palabra (errores ortográficos), o a nivel de la palabra y la oración (gramaticales o coherencia textual) (FERREIRA Y ELEJALDE, 2017).

2.6. Anotación del corpus

Siguiendo los procedimientos metodológicos desarrollados en los Proyectos FONDECYT señalados anteriormente, en cuanto a las anotaciones taxonómicas e implementación en el software UAM Corpus, se implementó una tipología

8. Complemento Indirecto

9. Corpus de referencia creado y administrado por la Real Academia Española y disponible en: <http://www.rae.es/recursos/banco-de-datos/corpes-xxi>. Este cuenta con más de con más 237.678 formas en su última versión 0.83 (1 de junio de 2016). Sus datos son provenientes de textos escritos y orales producidos en España, América, Filipinas y Guinea Ecuatorial entre los años 2001 y 2012 en una amplia gama de géneros y tipologías textuales.

específica para el estudio. A este programa se incorporaron los textos de los sub-corpus en formato .txt con codificación UTF-8, los que fueron etiquetados de acuerdo a los atributos identificados en dicho esquema (aceptabilidad, tipo y subtipo sintáctico-semántico), seleccionando los segmentos textuales en donde aparecían las unidades estudiadas. Asimismo, se ingresaron los datos de competencia de los sujetos. Posteriormente, mediante el uso de las aplicaciones del software fue posible recuperar informes específicos mediante el ingreso de parámetros de búsqueda, e informes estadísticos descriptivos e inferenciales.

3. ANÁLISIS, RESULTADOS Y DISCUSIÓN

3.1. ¿Cómo difiere la frecuencia y variedad de ColGram v+p de ambos niveles de competencia de estudiantes extranjeros con respecto a la producción de los HN?

Como se observa en la tabla 2, al analizar los usos de ColGram v+p en los tres sub-corpus, en el nivel A2 se identificaron 363 ocurrencias, con 122 colocaciones distintas entre ellas; mientras que en el nivel B1 se detectaron 422 ocurrencias, con 140 colocaciones distintas. Por su parte, los HN produjeron un total de 444 colocaciones, con una cifra de 267 unidades distintas.

Tabla 2: Densidad y variación colocacional según sub-sub-corpus

Índice	A2	B1	HN
Tamaño sub-corpus (n° de palabras)	24790	24789	24313
N° total de usos colocaciones	363	422	444
N° usos colocaciones /1000 palabras	14,64	17,02	18,26
N° colocaciones distintas	122	140	267
Proporción colocaciones distintas por usos	0,34	0,33	0,60

Basados en procedimientos utilizados en estudios que analizan colocaciones léxicas (OROL Y ALONSO, 2013; VINCZE, 2015), se calculó la frecuencia relativa de ColGram v+p respecto del total de palabras contenidas en cada sub-corpus, obteniéndose 14,64, 17,02, y 18,26 casos por mil palabras para los grupos A2, B1, y NH, respectivamente; lo que equivale al índice de densidad de ColGram v+p de cada sub-corpus. Luego, para comparar estos índices entre aprendientes y HN, se aplicó el equivalente de la prueba z en el programa estadístico R ('prop test', R CORE TEAM, 2013), nivel alfa de 0.05; en donde se obtuvo una diferencia estadística significativa entre las proporciones de los niveles A2 y los HN ($z=3,153$;

$p=0,002$); no así, entre el grupo B1 y los HN ($z=1,042$; $p=0,297$). Con estos datos, constatamos que el nivel A2 produjo un menor número de ColGram v+p que los HN; a diferencia del nivel B1 que no difiere en cuanto al total de unidades producidas por el grupo de HN. Estos hallazgos contradicen lo reportado por Rica (2010) sobre el uso de diversos tipos de ColGram en inglés como L2, en cuanto a que los hablantes no nativos utilizan un número mayor de ColGram que los HN dentro del campo léxico de los verbos de comunicación y procesos mentales.

Enseguida, para determinar si existían diferencias respecto de la cantidad de ColGram v+p distintas producidas por los aprendientes y HN se calculó el índice de variación colocacional (OROL Y ALONSO, 2013; VINCZE, 2015); es decir, la proporción de ColGram v+p distintas versus el número total de colocaciones producidas en cada grupo, obteniéndose las siguientes proporciones: A2=0,34, B1=0,33, y HN=0,60. Al comparar las proporciones de estos índices, se observó que existían diferencias estadísticas significativas entre los HN y el nivel A2 ($z=7.502$; $p=0,000$), y entre los HN y el grupo B1 ($z=7.945$; $p=0,000$).

Los antecedentes de los dos párrafos anteriores sugieren, que, aunque la frecuencia de ColGram v+p tiende a aumentar sustantivamente en el nivel B1, este incremento no va necesariamente acompañado de un aumento en el repertorio de colocaciones utilizadas. De acuerdo a ello, la producción de los HN estaría caracterizada por una mayor variación colocacional que la de los aprendientes de ambos niveles de competencia. Al respecto, aparecen ciertas líneas similares con lo planteado por Domínguez (2014) en su estudio en ELE en aprendientes intermedios de L1 valón y flamenco, quien concluye que el número de ColGram en general tendería a ser mayor que el de las colocaciones léxicas; con la observación que es en el grupo de colocaciones léxicas en donde se observa una mayor cifra de colocaciones distintas. Estos datos apuntarían a que el rasgo gramatical de las ColGram v+p (las preposiciones como una clase de palabras cerradas en la lengua) establece naturalmente restricciones en términos de las posibilidades de combinatoria, lo que limitaría el aumento potencial de ColGram distintas usadas en los escritos con un mejoramiento en el nivel de competencia, no así en su frecuencia.

Para profundizar en la naturaleza de estas tendencias, se examinaron las diez mayores frecuencias para determinar si algunas colocaciones específicas concentraban parte importante del total de las ColGram v+p de cada grupo y si estas diferían entre ellos; La estimación del uso repetido de algunas colocaciones en particular puede dar indicio de la caracterización del uso de las ColGram v+p de los aprendientes (FERREIRA Y ELEJALDE, 2017).

Los datos presentados en la tabla 3 revelan que al comparar las 10 mayores frecuencias colocacionales entre los tres grupos, encontramos que estas representan el 28,7% (A2), 30,8% (B1) y 20,7% (HN) de las frecuencias acumuladas dentro de cada sub-corpus. Así, también se obtiene la ocurrencia de 2 colocaciones en común entre los tres grupos, no obstante, con frecuencias y posiciones distintas en el ordenamiento dentro de ellos: *ir a*, con la mayor frecuencia en A2 (13,2%) y en B1 (6,2%); y la novena mayor frecuencia en los HN (1,1%); mientras que *ayudar a + infinitivo* obtiene la cuarta mayor frecuencia en A2 (2,2%), y la quinta mayor frecuencia en los grupos B1 (2,8%) y en los HN (2.0%).

Tabla 3: Frecuencia y porcentaje acumulado de colocaciones más usadas según sub-corpus.

A2 (N 363)				
Posición	ColGram v+p	Frecuencia absoluta	%	% acumulado
1	<i>ir a</i>	48	13,2	13,2
2	<i>viajar a</i>	9	2,5	15,7
3	<i>tratarse de (asunto)</i>	8	2,2	17,9
4	<i>ayudar a (+infinitivo)</i>	8	2,2	20,1
5	<i>compartir con</i>	6	1,7	21,8
6	<i>hablar sobre</i>	5	1,4	23,1
7	<i>apoyar a (+infinitivo)</i>	5	1,4	24,5
8	<i>quedarse en</i>	5	1,4	25,9
9	<i>cantar sobre</i>	5	1,4	27,3
10	<i>escribir sobre</i>	5	1,4	28,7

B1 (N 422)				
Posición	ColGram v+p	Frecuencia absoluta	%	% acumulado
1	<i>ir a</i>	26	6,2	6,2
2	<i>tratar de (asunto)</i>	22	5,2	11,4
3	<i>hablar sobre</i>	17	4,0	15,4
4	<i>escribir sobre</i>	16	3,8	19,2
5	<i>ayudar a (+infinitivo)</i>	12	2,8	22,0
6	<i>tratarse de (asunto)</i>	8	1,9	23,9
7	<i>hacer de (materia)</i>	8	1,9	25,8
8	<i>hablar de</i>	8	1,9	27,7
9	<i>saber sobre</i>	6	1,4	29,1
10	<i>interesarse en</i>	6	1,4	30,6

HN (N 444)				
Posición	ColGram v+p	Frecuencia absoluta	%	% acumulado
1	llegar a	20	4,5	4,5
2	pensar en	13	2,9	7,4
3	hablar de	11	2,5	9,9
4	llevar a	9	2,0	11,9
5	ayudar a (+infinitivo)	9	2,0	14,0
6	entrar a	8	1,8	15,8
7	convertirse en	7	1,6	17,3
8	sacar de	5	1,1	18,5
9	ir a	5	1,1	19,6
10	pasar por	5	1,1	20,7

Respecto de *ir a*, en la comparación entre las proporciones de las frecuencias en los niveles A2 y HN, se obtiene una diferencia estadística significativa entre ellas ($z=6901$; $p=0,000$), con mayor frecuencia en el nivel A2; y entre B1 y HN ($z=3,986$; $p=0,000$), con mayor frecuencia en el nivel B1. La alta frecuencia de dicha colocación en los aprendientes resultaría de la elicitación a partir de las temáticas de tareas específicas (temas 4, 9, 10, y 11) asignadas a los aprendientes, por ejemplo, el caso del tema 4 *Experiencias en las vacaciones* claramente induciría a los aprendientes a expresar el significado de dirección y destino mediante la producción de la colocación *ir a* (tabla 1). Por otra parte, no se obtienen diferencias estadísticas entre las proporciones de las frecuencias de la colocación *ayudar a + infinitivo* entre los aprendientes y los HN. Adicionalmente a las dos colocaciones señaladas, en los 2 grupos de aprendientes aparecen las colocaciones *tratarse de* (A2=2,2%; B1=1,9%) y *hablar sobre* (A2=1,4%; B1=4.0), mientras que entre los grupo B1 y HN aparece en común la colocación *hablar de* (B1=1,9%; HN=2,5%).

Complementariamente a lo anterior, los datos del ordenamiento de las ColGram v+p más frecuentes en los tres sub-corpus apuntan en la dirección señalada anteriormente: algunas colocaciones específicas tienden a concentrar un número importante de las ocurrencias en los sub-corpus A2 y B1. Por ejemplo, las cinco colocaciones más utilizadas en ambos grupos equivalen a aproximadamente el 22% de las ocurrencias en ambos sub-corpus de aprendientes; a su vez, las diez colocaciones con las mayores frecuencias en ambos niveles alcanzan al 30% aproximadamente del total de colocaciones producidas en ambos niveles. En contraste, en el caso de los HN las cinco y diez mayores frecuencias solo explican el 14,0% y 20,7% de las ocurrencias en este sub-corpus, respectivamente.

En definitiva, los datos de los aprendientes muestran la sobrerrepresentación de algunas colocaciones en particular, que explicarían, por una parte, los datos de alta densidad colocacional del nivel B1, el cual no se diferencia del grupo de HN en la frecuencia, y por otra parte, la variación colocacional menor en ambos sub-corpus de aprendientes comparado con los HN. Por ejemplo, considerando solo las frecuencias iguales o mayores a 10 ocurrencias en cada sub-corpus, en el nivel A2 tenemos solo a la colocación *ir a*, la cual explica el 13,2% de todas las ocurrencias colocacionales en dicho nivel, mientras que en el nivel B1 las colocaciones *ir a*, *tratar de (asunto)*, *hablar sobre*, *escribir sobre* y *ayudar a (+ infinitivo)* en conjunto dan cuenta del 22% de las ocurrencias. Por su parte, en el sub-corpus de HN, las colocaciones con frecuencias mayores a 10 ocurrencias solo explican el 9,9% de los casos. Estos hallazgos parecieran indicar una cierta similitud respecto de las tendencias del estudio de colocaciones léxicas en inglés como L2 de Granger y Bersten (2014); al comparar la frecuencia de los aprendientes intermedios versus los avanzados y los HN, los aprendientes de menor competencia muestran un mayor uso de colocaciones de mayor frecuencia en la lengua, y, por el contrario, muestran un uso infrarrepresentado de colocaciones menos frecuentes.

Una variable adicional que podría aportar a explicar los datos de variación colocacional entre los sub-corpus es si en ellos hay bastante diversidad temática para gatillar el uso de las colocaciones. En este sentido, el resultado de mayor variación del grupo de HN que de los aprendientes derivaría de la sub-representación de algunas áreas temáticas en los sub-corpus de aprendientes (SÁNCHEZ, 2015c), lo que estimularía una menor producción de colocaciones distintas. En concreto, los aprendientes versaron sus escritos sobre 16 (B1) y 20 (A2) temas distintos, los que eliciarían una menor variación colocacional comparado con los temas elegidos libremente por los HN en las tareas 2 y 3.

3.2. ¿Cómo difieren los aprendientes extranjeros de los HN en la frecuencia de usos correctos de ColGram v+p según los patrones sintáctico-semánticos que describen a estas unidades?

Para examinar la frecuencia de acuerdo al patrón sintáctico gramatical, el análisis de los usos correctos e incorrectos de las ColGram v+p se realizó separadamente. En primer lugar, identificamos el número de usos colocacionales

correctos versus los usos incorrectos del total de ColGram v+p producidas por cada grupo. El número de usos correctos correspondió a 210 (57,9%) en A2, 280 (66,4%) en B1, y 434 (97,8%) en HN. En estos tres casos, al comparar las proporciones de usos correctos con los usos incorrectos, se obtuvo una diferencia estadística significativa entre las dos condiciones dentro de cada grupo (A2: $z=4,231$, $p=0,000$; B1: $z=9,500$, $p=0,000$; HN: $z=2,846$, $p=0,000$). Con estos datos se evidencia que las proporciones de los usos correctos superan a las de los incorrectos en los tres grupos, coincidiendo con lo reportado por Bahardoust y Moeini (2012) en inglés como L2, respecto de que la proporción de distintos tipos de ColGram producidas correctamente supera a los usos incorrectos.

Para responder a la interrogante de cómo difieren los aprendientes de los HN en los usos correctos, a lo largo de los cuatro niveles de análisis definidos por la taxonomía colocacional, se empleó la prueba de Chi Cuadrado¹⁰. El programa UAM Corpus Tool incorpora entre las herramientas de análisis el cálculo de Chi Cuadrado, lo que permite emitir informes indicando el nivel de significancia estadística entre las diferencias de frecuencia expresada mediante la siguiente nomenclatura: (1) +, significativo en el nivel del 90%, 10% probabilidad de error, (2) ++, significativo en el 95%, 5% probabilidad de error, y (3) (+++), significativo en el 98%, con 2% probabilidad de error. Como se aprecia en la tabla 5, al comparar los datos entre aprendientes y HN mediante la aplicación de este estadístico, como es esperable, se observan diferencias estadísticas significativas entre ellos, en donde los HN obtienen un mayor número de ColGram v+p correctas que ambos niveles de aprendientes, y una menor tasa de error.

Tabla 4: Frecuencia y porcentaje de usos correctos e incorrectos de los sub-corpus de aprendientes y HN.

	A2 (N=363)		HN (N=444)		Chisqu	Signif.
	Frecuencia	%	Frecuencia	%		
Usos correctos	210	57,9	434	97,8	197,2	+++
Usos incorrectos	153	42,1	10	2,2	197,2	+++

	B1 (N=422)		HN (N=444)		Chisqu	Signif.
	Frecuencia	%	Frecuencia	%		
Usos correctos	280	66,4	434	97,8	147,4	+++
Usos incorrectos	142	33,6	10	2,2	147,4	+++

10. Estadístico utilizado comúnmente en estudios lingüísticos (O'Donnell 2008) para comparar datos independientes de naturaleza nominal, como es la frecuencia, con el objetivo de determinar si dos variables funcionan independientemente entre sí. En este caso, la hipótesis nula representa la independencia entre los datos (Choura, 2017).

Como se observa en la tabla 4, de los cuatro patrones sintáctico-semánticos identificados en la taxonomía colocacional, en los tres sub-corpus se obtiene un mayor número de colocaciones correspondientes a los patrones 1 (*Verbo + 2 ó más preposiciones alterantes con un significado similar*): (A2=71,4%; B1=72,5%; HN=44,2%), y 4 (*Verbo + CPrep que puede aparecer tácito*): (A2=20,0%; B1=17,5%; HN=33,9%). Por otra parte, los patrones 2 y 3 muestran las menores frecuencias colocacionales en los tres grupos de participantes. También se puede apreciar que de los cuatro patrones sintáctico-semánticos es el grupo de HN el que presenta las mayores frecuencias numéricas de los tres sub-corpus en los patrones 2, 3 y 4, mientras que el grupo B1 es el que obtiene la mayor frecuencia en el patrón 1. En contraste, el grupo A2 es el que tiene las menores frecuencias en los cuatro patrones.

Tabla 5: Frecuencia y porcentaje de usos correctos según patrón sintáctico-semántico y por sub-corpus.

Patrón sintáctico semántico	A2		B1		HN	
	Frecuencia	%	Frecuencia	%	Frecuencia	%
Patrón 1	150	71,4	203	72,5	192	44,2
Patrón 2	5	2,4	15	5,4	27	6,2
Patrón 3	13	6,2	13	4,6	68	15,7
Patrón 4	42	20,0	49	17,5	147	33,9
	210	100	280	100	434	100

Luego, en la comparación específica entre niveles de competencia y HN, en primer lugar, evidenciamos en la tabla 6 que los HN superan la frecuencia del grupo A2 en los patrones 2, 3 y 4 (los tres con alta significancia); mientras que en el caso del patrón 1, aunque la frecuencia del grupo HN supera numéricamente al grupo A2, no se observan diferencias estadísticas de significancia entre los dos sub-corpus.

Tabla 6: Comparación de frecuencia de usos correctos por patrón sintáctico-semántico entre niveles A2 y HN.

Patrón sintáctico semántico	A2 (N=363)		HN (N=444)		Chisqu	Signif.
	Frecuencia	%	Frecuencia	%		
Patrón 1	150	41,3	192	43,2	0,30	
Patrón 2	5	1,4	27	6,1	11,60	+++
Patrón 3	13	3,6	68	15,3	30,45	+++
Patrón 4	42	11,6	147	33,1	51,66	+++
Total usos correctos	210		434			

En segundo lugar, como se observa en la tabla 7, al comparar las frecuencias entre el grupo B1 y los HN, se obtiene que los HN superan con alta significancia estadística al nivel B1 en los patrones 3 y 4, mientras que en el caso del patrón 2, los HN sobrepasan al grupo B1 con un nivel de significancia estadística marginal. En cuanto al patrón 1, no se aprecian diferencias estadísticas significativas entre los HN y el nivel B1. Estos datos y los del párrafo anterior muestran que la producción de ColGram v+p correctas de los HN supera en frecuencia, como es esperable, a ambos grupos de aprendientes, excepto en el patrón 1.

Tabla 7: Comparación de frecuencia de usos correctos por patrón sintáctico-semántico entre niveles B1 y HN.

Patrón sintáctico semántico	B1 (N=422)		HN (N=444)		Chisqu	Signif.
	Frecuencia	%	Frecuencia	%		
Patrón 1	203	48,1	192	43,2	2,06	
Patrón 2	15	3,6	27	6,08	2,99	+
Patrón 3	13	3,1	68	15,3	38,20	+++
Patrón 4	49	11,6	147	33,1	57,10	+++
Total usos correctos	280		434			

En un análisis más detallado por subtipo dentro de los patrones sintáctico-semánticos, como se presenta en la tabla 8, se observa que en los tres grupos de la muestra las mayores frecuencias relativas se sitúan en los subtipos de los patrones 1 y 4, mientras que las menores ocurren en los subtipos de los patrones 2 y 3, con proporciones menores al 5%, exceptuando al subtipo 3.1 en el grupo de HN, que alcanza el 6,9%.

Tabla 8: Frecuencia y porcentaje de usos correctos por subtipo sintáctico-semántico según sub-corpus.

Patrón sintáctico semántico	Subtipo	A2 (N=363)		B1 (N=422)		HN (N=444)	
		Frecuencia	%	Frecuencia	%	Frecuencia	%
Patrón 1	1.1	57	27,1	134	47,9	75	17,3
	1.2	83	39,5	52	18,6	74	17,1
	1.3	10	4,8	17	6,1	43	9,9
Patrón 2	2.1	2	1,0	3	1,1	9	2,1
	2.2	1	0,5	11	3,9	16	3,7
	2.3	2	1,0	1	0,4	2	0,5
Patrón 3	3.1	4	1,9	4	1,4	30	6,9
	3.2	5	2,4	4	1,4	14	3,2
	3.3	1	0,5	4	1,4	10	2,3
	3.4	3	1,4	1	0,4	14	3,2
Patrón 4	4.1	21	10,0	25	8,9	35	8,1
	4.2	9	4,3	17	6,1	38	8,8
	4.3	12	5,7	7	2,5	74	17,1
Total usos correctos		210	100	280	100	434	100

Por otra parte, en los sub-corpus de aprendientes sobresalen algunas frecuencias de subtipos individuales que concentran una cifra relativamente alta del total de colocaciones producidas correctamente en cada sub-corpus. El nivel A2 contiene dos subtipos (1.1=27,1%; 1.2=39,5) que en conjunto explican el 66,6% de los casos detectados; en forma similar, estos mismos dos subtipos concentran el 66,4% en el nivel B1 (1.1=47,9%; 1.2=18,6). Asimismo, en el caso del sub-corpus de HN destacan estos mismos dos subtipos, pero con una representación porcentual considerablemente menor, equivalente al 34,4% del total de ocurrencias (1.1=17,3%; 1.2=17,1), además del subtipo 4.3 con un 17,1% del total de colocaciones correctas.

Al realizar la comparación de los usos correctos de ColGram v+p entre el nivel A2 y el grupo de HN, como se muestra en la tabla 9, se detectó que no existían diferencias de significancia estadística entre ambos sub-corpus en los subtipos 1.1,

2.3, y 3.1. En los otros 10 subtipos, las frecuencias entre los dos grupos difieren entre sí con diferentes grados de significancia estadística: alta en los subtipos 1.3, 2.2, 3.1, 3.3, 4.2, 4.3; moderada en los subtipos 1.2, y 3.4; y leve en el caso de los subtipos 2.1, y 3.2. En todos estos casos la frecuencia es mayor en el grupo de HN que en el nivel A2, con excepción del grupo 1.2, en donde el grupo de aprendientes supera a los HN.

Tabla 9: Comparación de frecuencia de usos correctos por subtipo sintáctico-semántico entre niveles A2 y HN.

Patrón sintáctico semántico	Subtipo	A2 (N=363)		HN (N=444)		Chisqu	Signif.
		Frecuencia	%	Frecuencia	%		
Patrón 1	1.1	57	15,7	75	16,9	0,21	
	1.2	83	22,9	74	16,7	4,99	++
	1.3	10	2,8	43	9,7	15,63	+++
Patrón 2	2.1	2	0,6	9	2,0	3,24	+
	2.2	1	0,3	16	3,6	10,73	+++
	2.3	2	0,6	2	0,5	0,04	
Patrón 3	3.1	4	1,1	30	6,8	15,83	+++
	3.2	5	1,4	14	3,2	2,74	+
	3.3	1	0,3	10	2,3	5,81	+++
	3.4	3	0,8	14	3,2	5,24	++
Patrón 4	4.1	21	5,8	35	7,9	1,36	
	4.2	9	2,5	38	8,6	13,46	+++
	4.3	12	3,3	74	16,7	37,45	+++

Posteriormente, al comparar la producción del grupo B1 con la de los HN en la tabla 1, se encontró que la frecuencia no difería estadísticamente en 4 cuatro subtipos: 2.2, 2.3, 3.3, y 4.1. Por otra parte, se identificaron diferencias estadísticas con diferentes grados de significancia entre los dos grupos: de tipo alta (subtipos 1.1, 1.3, 3.1, 3.4, 4.2, y 4.3), moderada en el caso del subtipo 3.2, y leve en el caso de los subtipos 1.2, y 2.1. En todos estos casos el grupo de HN supera al nivel B1, exceptuando al subtipo 1.1, en donde el nivel B1 supera la frecuencia del grupo de HN.

Tabla 10: Comparación de frecuencia de usos correctos por subtipo sintáctico-semántico entre niveles B1 y HN.

Patrón sintáctico semántico	Subtipo	B1 (N=422)		HN (N=444)		Chisqu	Signif.
		Frecuencia	%	Frecuencia	%		
Patrón 1	1.1	134	31,8	75	16,9	26,10	+++
	1.2	52	12,3	74	16,7	3,28	+
	1.3	17	4,0	43	9,7	10,74	+++
Patrón 2	2.1	3	0,7	9	2,0	2,74	+
	2.2	11	2,6	16	3,6	0,71	
	2.3	1	2,0	2	0,5	0,29	
Patrón 3	3.1	4	1,0	30	6,8	19,36	+++
	3.2	4	1,0	14	3,2	5,17	++
	3.3	4	1,0	10	2,3	2,32	
	3.4	1	0,2	14	3,2	10,81	+++
Patrón 4	4.1	25	5,9	35	7,9	1,29	
	4.2	17	4,0	38	8,6	7,47	+++
	4.3	7	1,7	74	16,7	57,49	+++

Luego, para indagar en mayor profundidad la naturaleza de las tendencias observadas en función de la frecuencia obtenida en los patrones sintáctico-semánticos generales, se analizaron en mayor detalle los subtipos incluidos dentro de estos patrones. En primer lugar, en el subtipo 1.1 (*Verbo + (CD) + CPrep con preposiciones alternantes*) la frecuencia del nivel B1 (31,8%) supera en más del doble la proporción de las colocaciones correctas del nivel A2 (15,7%). Al respecto, un aspecto que podría dar cuenta de esta diferencia sería el uso recurrente de colocaciones específicas dentro de los sub-corpus. La tabla 11 muestra separadamente la frecuencia de colocaciones que tienen tres o más ocurrencias del subtipo 1.1 en los tres sub-corpus. Como se puede observar, la mayor frecuencia del grupo B1 comparada con el nivel A2 puede explicarse a partir los usos sobrerrepresentados de *escribir sobre* (16 ocurrencias), *hablar sobre* (17 ocurrencias), y *tratar de algo* (22 ocurrencias), correspondientes al 77,6% de las frecuencias del subtipo 1.1. Estas resultan ser muy superiores a las producidas en el nivel A2 (5, 3, y 3 ocurrencias, respectivamente). Por su parte, en los HN se cuentan solo 5 colocaciones que igualan o superan las 3 ocurrencias dentro del sub-corpus, comparado con las 14 colocaciones del nivel B1, y con las 7 colocaciones del nivel A2. Estos datos en general apuntan al uso repetido de varias colocaciones del subtipo 1.1 en los grupos de aprendientes, con mayor notoriedad en el nivel B1, que explicaría la alta frecuencia numérica de colocaciones comparado con el grupo de HN y que contribuye a que no se observe

una diferencia estadística entre las frecuencias de los aprendientes y los HN en el patrón 1.

Tabla 11: Frecuencia de ColGram más frecuentes (3 ocurrencias o más) en el subtipo sintáctico-semántico 1.1 según sub-corpus

Patrón 1. Subtipo 1.1 ColGram v+p	Frecuencia		
	A2 (N=57)	B1 (N=134)	HN (N=75)
aprender de (asunto)	3	4	-
aprender sobre (asunto)	5	4	-
asociar a (asunto)	-	-	3
cantar de (asunto)	5	3	-
conocer de (asunto)	-	3	-
enseñar sobre (asunto)	-	3	-
enviar a (destino)	-	3	-
escribir sobre (asunto)	5	16	-
escuchar en (medio)	-	4	-
hablar de (asunto)	-	8	11
hablar sobre (asunto)	5	17	-
hacer de (materia)	-	6	-
ligar a (asunto)	-	-	3
llevar a (destino)	-	-	9
saber sobre (asunto)	3	6	-
sacar de (procedencia)	-	-	5
tratar de (asunto)	3	22	-
tratar sobre (asunto)	-	5	-
Total	29 (50,9%)	104 (77,6%)	31 (41,5%)

En cuanto al subtipo 1.2, siguiendo el análisis anterior, tenemos una situación similar, pero que se manifiesta de forma inversa considerando el nivel de competencia en la lengua, como se ilustra en la tabla 12. El grupo A2 supera la frecuencia total del nivel B1 en este subtipo debido a la alta frecuencia de dos colocaciones en particular: *ir a* y *viajar a*, con 53 ocurrencias de un total de 83 en este subtipo en el nivel A2. En definitiva, las frecuencias colocacionales iguales o mayores a 3 ocurrencias en el nivel A2 explican el 81,9% de los casos, comparado con los porcentajes considerablemente menores en el grupo B1 (65,4%) y NH (54,1%). De esta forma, al uso repetido de algunas colocaciones, en particular, *ir a* en el nivel A2, aparece como un rasgo que explica la alta cifra en la producción

colocacional del subtipo 1.2 de los aprendientes, contribuyendo también a que no exista diferencia entre la frecuencia de los aprendientes de nivel A2 y los HN.

Tabla 12: Frecuencia de las ColGram v+p más frecuentes (3 ocurrencias o más) en el subtipo sintáctico-semántico 1.2 según la competencia

Patrón 1. Subtipo 1.2 ColGram v+p	Frecuencia		
	A2 (N=83)	B1 (N=52)	HN (N=74)
ir a (destino)	48	26	5
entrar a (destino)	-	-	8
entrar en (destino)	-	-	4
llegar a (destino)	-	-	20
subir a (destino)	-	-	3
regresar a (destino)	3	-	-
venir a (destino)	-	4	-
venir de (procedencia)	4	-	-
viajar a (destino)	9	4	-
volver a (destino)	4	-	-
Total	68 (81,9%)	34 (65,4%)	40 (54,1%)

En el patrón 2, se observa una alta diferencia estadística entre el grupo A2 y HN, y leve entre el grupo B1 y HN. En este caso, se puede observar que con el mejoramiento de la competencia habría un acercamiento a la producción colocacional nativa. Por ejemplo, dentro del subtipo 2.2 se incluye potencialmente un grupo reducido de verbos, tales como, *alentar a*, *animar a*, *apoyar a*, *impulsar a*, *ayudar a*, *influenciar a*, y *presionar a*, los que se enmarcan dentro de la estructura *verbo + (CD) + a + infinitivo*. Intuimos que la diferencia en frecuencia de usos correctos de este subtipo entre el nivel A2 (1 ocurrencia) y la del nivel B1 (11 ocurrencias) podría originarse en la complejidad que presentarían estas colocaciones para los aprendientes del nivel A2, la que se caracteriza por la posición indirecta del CPrep respecto del predicado, y porque además exige una categoría gramatical específica, un infinitivo, como término de preposición del sintagma preposicional (VAISBERG, 2013). En cuanto al grupo B1, este no muestra diferencias al grupo de HN, lo que se explicaría por el avance lingüístico de los aprendientes.

Otra tendencia destacada en el de los subtipos de colocaciones que incluyen usos pronominales de verbos. Observamos que se obtiene una diferencia estadística significativa entre aprendientes y HN, en donde los HN superan a ambos niveles en los subtipos 1.3 (*acercarse a*, *inscribirse a*, etc.), 3.4 (*adaptarse a*, *fijarse en*, etc.), y 4.3

(*caracterizarse por, reírse de*, etc.), y además en el nivel A2 en el subtipo 3.3 (*encontrarse con, desempeñarse en*, etc.). Estos datos apuntarían a que tanto la producción de ambos niveles tendería a caracterizarse por usos infrarrepresentados de ColGram v+p que incluyen la producción de pronominales, por lo que estimamos que este tipo de usos colocacionales constituyen una fuente de dificultad para los aprendientes de ambos grupos.

Finalmente, en el patrón 4.2 (*Verbo no pronominalizado + CPrep directo que puede aparecer implícito*, por ejemplo, *asistir a, dudar de*, etc.), también es posible observar un comportamiento similar a los de los verbos pronominalizados. Los aprendientes muestran usos infrarrepresentados en este tipo de colocaciones, comparado con la producción de los HN. Lo anterior también va en línea de explicar el rasgo de menor variedad que caracteriza la producción de los aprendientes.

CONCLUSIONES

El objetivo de este estudio fue examinar el uso de las ColGram v+p 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 de HN. Los datos apuntan a rasgos de naturalidad restringida en los aprendientes. Esto se deduce a partir de la menor frecuencia de uso de los aprendientes, lo que mejora en el nivel B1; no obstante, la variedad limitada en la producción de estas unidades en los aprendientes es el rasgo que aparece con mayor notoriedad puesto que persiste en el nivel B1, diferenciándose del uso de HN. En términos prácticos, esto se presenta como el uso recurrente de algunas unidades en particular, sobresaliendo los usos sobrerrepresentados de *ir a* en ambos grupos de aprendientes, y de *tratar de* (asunto) en el grupo B1, junto a otras unidades colocacionales que concentran gran parte de las frecuencias en los sub-corpus de aprendientes.

En relación con la identificación de los usos correctos e incorrectos, se evidenció que ambos niveles de aprendientes difieren de los HN en la proporción de usos correctos y tasa de error; produciendo los aprendientes menos usos adecuados, y por ende mayor cantidad de usos incorrectos desde el punto de vista de la combinación v+p esperada.

Por otra parte, en cuanto a los 4 patrones sintáctico-semánticos identificados en la tipología colocacional, se aprecia que en general los aprendientes de ambos niveles muestran similitud con los HN en la proporción de usos correctos en el patrón colocacional 1 (*verbo + 2 ó más preposiciones alternantes con un significado similar*),

el que constituye el grupo de mayor frecuencia en cada uno de los tres sub-corpus analizados. Por el contrario, los aprendientes difieren en los patrones 2, 3, y 4 respecto de los HN, con una menor frecuencia de usos adecuados en ellos.

En cuanto al análisis detallado de los subtipos colocacionales específicos, se observan usos sobrerrepresentados, que ayudan a explicar a nivel global que no exista diferencia en la frecuencia colocacional, pero sí diferencias en la variedad colocacional, entre aprendientes y nativos. Así tenemos, en el subtipo 1.1 en el grupo B1 usos sobrerrepresentados de unidades tales como: *escribir sobre, hablar sobre y tratar de algo*; mientras que en el subtipo 1.2 sobresalen los usos sobrerrepresentados de *ir a y viajar en* ambos grupos de aprendientes, con mayor magnitud en el grupo A2. Por otra parte, como parte de las diferencias entre aprendientes y HN, obtenemos en los aprendientes una infrarrepresentación de unidades en los subtipos colocacionales que incluyen usos pronominales de los verbos (1.3, 3.3, 3.4 y 4.3), y en el grupo 4.2.

Este trabajo responde a la necesidad de lograr descripciones más completas de la interlengua colectiva de aprendientes anglófonos de ELE (SÁNCHEZ, 2015a), y de proporcionar datos de mayor especificidad en el uso de ColGram v+p (HENRIKSEN, 2013), aplicando una clasificación detallada de patrones y subtipos en base a un criterio sintáctico-semántico. Como limitaciones del trabajo, se puede mencionar que la evidencia sobre la que se basa el estudio corresponde a datos de la actuación de los estudiantes, y por ello, no nos permite realizar estimaciones acerca de aquellas unidades colocacionales no producidas en las tareas que conforman los corpus escritos. Así también, estos datos no entregan información relativa a la comprensión o el procesamiento de las unidades estudiadas. Finalmente, este trabajo no profundiza en un análisis detallado de las ColGram v+p producidas en forma incorrecta, respecto de su descripción y explicación, etapas posteriores contempladas en la investigación mayor en la que se enmarca el estudio aquí presentado.

REFERENCIAS

- BAHARDOUST, M.; MOEINI, M. (2012). Lexical and grammatical collocations in writing production of EFL learners. *The Journal of Applied Linguistics*. v. 5, n° 1, pp. 61-86.
- BENSON, M.; BENSON, E.; ILSON, R. (1986): *The BBI combinatory dictionary of English. A guide to word combinations*. Amsterdam: John Benjamins Publishing Company.
- BOSQUE, L. (2001). Sobre el concepto de colocación y sus límites. *Lingüística española actual*. v. 13, n° 1, pp. 9-41.

- BUYSE, K. (2003), Colocaciones léxicas: pistas y trampas. *Mosaico*. n° 10, pp. 10-18.
- CHEIKH-KHAMIS, F. (2013). *Lingüística cognitiva aplicada a la enseñanza del léxico en segundas lenguas. Las colocaciones en clase de ELE*. Trabajo Fin de Máster. Máster en Ciencia del lenguaje y filología hispánica. UNED.
- CONSEJO DE EUROPA. (2001). Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación. Consejo de Europa, Estrasburgo. En línea. Disponible en: http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/: Acceso el: 1/junio/2018.
- DOMÍNGUEZ, U. (2014). Las colocaciones en un corpus de aprendices valón y flamenco. In: Morimoto, Y.; Pavón, M.; Santamaría, R. (eds.) *Asociación para la Enseñanza del Español como Lengua Extranjera, La enseñanza de ELE centrada en el alumno. Congreso Internacional 25, 2014*, pp. 977-986.
- FERNÁNDEZ, G. (2014). Enseñanza-Aprendizaje de las colocaciones en el nivel inicial (A1-A2). *Marcoele*. n° 19, pp. 1-19.
- FERREIRA, A. (2014). *Proyecto de investigación FONDECYT - CONICYT N°1440651: El feedback correctivo escrito directo e indirecto en la adquisición y aprendizaje del español como lengua extranjera*. Universidad de Concepción, Chile 2014-2017.
- FERREIRA, A. (2014-2018). *Corpus de Aprendices de Español Lengua Extranjera, CAELE*, Universidad de Concepción, Chile.
- FERREIRA, A. (2016). *Prueba de Multinivel con Fines Específicos Académicos, Programa de Español como Lengua Extranjera ELE-UdeC*, Universidad de Concepción, Chile.
- FERREIRA, A. (2018). *Proyecto de investigación FONDECYT - CONICYT N°1180974: Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua*. Universidad de Concepción, Chile.
- FERREIRA, A.; ELEJALDE, J. (2017). Análisis de errores recurrentes en un Corpus de aprendices de español como lengua extranjera (Corpus CAELE). *Revista Brasileira de Lingüística Aplicada*. v. 17, n° 3, pp. 509-538.
- FERREIRA, A.; ELEJALDE, J.; VINE, A. (2014). Análisis de errores asistido por computador basado en un corpus de aprendientes de español como lengua extranjera. *Signos*. v. 46, n°86, pp. 385-411.

- FERREIRA, A.; LAFLEUR, N. (2015). Analyse et description des erreurs prépositionnelles les plus fréquentes en espagnol L2. *Linguística y Literatura*. v. 68, pp. 57-79.
- FERREIRA, A., OPORTUS, R. (2018). Procesamiento cognitivo del Feedback Correctivo Escrito indirecto en los errores preposicionales en ELE. *Boletín de Filología*. v. 53, n°1, pp. 83-108.
- FERREIRA, A., OPORTUS, R.; FUENTES, K. (2016). Tiempos de respuesta y feedback correctivo escrito en aprendientes de ELE. *Revue Française de Linguistique Appliquée*. v. 22, n° 2, pp. 109-122.
- GONZÁLEZ, F. (2006). Las colocaciones en la enseñanza del español de los Negocios. *Revista de Didáctica MarcoELE*. n° 2, pp. 61-86.
- GRANGER, S. (2002). A bird's-eye view of learner corpus research. In: Granger, S. et al. (Eds.), *Computer Corpora, Second Language Acquisition and Foreign Language Teaching*, Amsterdam: Benjamins. pp. 3-33.
- GRANGER, S. (2015). Contrastive interlanguage analysis: A reappraisal, *International Journal of Learner Corpus Research*. v. 1, n° 1, pp. 7-24.
- GRANGER, S.; BESTGEN, Y. (2014). The use of collocations by intermediate vs. advanced and non-native writers: a bigram based study. *International Review of Applied Linguistics in Language Teaching*. v. 52, n° 1, pp. 229-252.
- GRANGER, S.; PAQUOT, M. (2008). Disentangling the phraseological web. In: Granger S.; Meunier F., *Phraseology: An Interdisciplinary Perspective*. Amsterdam: Benjamins: pp. 27-49.
- GRANGER, S.; PAQUOT, M.; RAYSON, P. (2006). Extraction of multi-word units from EFL and native English corpora. The Phraseology of the verb 'make'. En Buhofer, H.; Burger, H. *Phraseology in Motion I: Methoden und Kritik. Akten der Internationalen Tagung zur Phraseologie*. Schneider Verlag Hohengehren: Baltmannsweiler, pp. 57-68.
- GYLLSTAD, H.; WOLTER, B. (2015). Collocational Processing in Light of the Phraseological Continuum Model: Does Semantic Transparency Matter? *Language Learning*. v. 66, n° 2, pp. 296-323.
- HENRIKSEN, B. (2013). Research on L2 learners' collocational competence and development – a progress report. En Bardel, C.; Lindquist, C.; Laufer, B. (eds.). *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis*. Eurosla Monographs Series 2. Amsterdam: Eurosal, pp 29-57.

- HIGUERAS M. (2004). *La enseñanza-aprendizaje de las colocaciones en el desarrollo de la competencia léxica en el español como lengua extranjera*. Tesis doctoral. Universidad Complutense de Madrid, Madrid.
- HOWARTH, P. (1996). *Phraseology in English Academic Writing*. Tübingen: Max Niemeyer Verlag.
- JIMÉNEZ, L. (2011). El estatus argumental del complemento de régimen prepositivo en español. *Signos Lingüísticos*. v. 7, n° 14, pp. 81-114.
- KOIKE, K. (2001). *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Alcalá: Universidad de Alcalá de Henares.
- LU, L. (2016). Colocaciones usadas en las pruebas de lectura DELE B1. *Círculo de Lingüística Aplicada a la Comunicación*. n° 66, pp. 210-243.
- O'DONNELL, M. (2008). The UAM CorpusTool: software for corpus annotation and exploration. Proceedings of the XXVI Congreso de AESLA, Almería, Spain, 3-5 April 2008.
- OROL, A.; ALONSO, M. (2013). A comparative study of collocations in a native corpus and a learner corpus of Spanish. *Procedia Social and Behavioral Sciences*. n°. 95, pp. 563-570.
- PÁEZ, G. (2004). Géneros periodísticos en manuales de estilo y de periodismo. *Revista científica de UCES*. 8(1). v. 8, n° 1, pp. 80-98.
- R CORE TEAM (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponible en: <http://www.R-project.org/> : Acceso el: 1/junio/2018.
- RAE. (2009). *Nueva gramática de la lengua española Manual*. Madrid: Espasa.
- RAE. (2018). Banco de datos (CORPES XXI), Corpus del Español del Siglo XXI. [En línea] Disponible en: <http://web.frl.es/CORPES/view/inicioExterno.view>. Acceso 14/05/2018.
- RICA, J. (2010). Colocaciones gramaticales en la producción escrita de estudiantes universitarios españoles. *Reduca (Filología)*. v. 2, n° 1, pp. 1-25.
- RODRÍGUEZ-FERNÁNDEZ, S.; CARLINI, R.; WANNER, L. (2015). Classification of grammatical errors in the writings of learners of Spanish. *Procesamiento del lenguaje natural*. n° 55, pp. 49-56.

- SÁNCHEZ, A. (2010). Apuntes sobre las colocaciones léxicas y el concepto de colocación. *Anuario de Estudios Filológicos*. n° 33, pp. 291-306.
- SÁNCHEZ, A. (2015a). Las funciones diagnóstica y evaluativa del análisis contrastivo de la interlengua del español basado en corpus. *Linred*. n° 13.
- SÁNCHEZ, A. (2015b). *El verbo dar en el español escrito de aprendientes de L1 inglés: estudio comparativo entre hablantes no nativos y hablantes nativos basado en corpus*. Tesis doctoral. Departamento de Filología y Lingüística General. Universidad de Extremadura.
- SÁNCHEZ, A. (2015c). Análisis contrastivo de interlengua y corpus de aprendientes: precisiones metodológicas. *Pragmalingüística*. n° 23, pp. 191-210.
- TRAVALIA, C. (2006). Las colocaciones gramaticales en español. *Anuario de estudios filológicos*. n° 29, pp. 279-295.
- VAISBERG, R. (2013) *Cinco nuevas funciones sintácticas en el español*. Trabajo de ascenso para optar al escalafón de Profesor Agregado. Universidad Católica Andrés Bello. Facultad de Humanidades y Educación.
- VINCZE, O. (2015). *Learning multiword expression from corpora and dictionaries*. Doctoral Thesis. Departamento de Galeno, Portugués, Francés e Linguística. Universidade da Coruña.

Recebido: 8/11/2018

Aceito: 16/07/2019

Publicado: 7/08/2019