

<https://doi.org/10.32735/S0718-2201202000050790>

179-193

COLOCACIONES GRAMATICALES VERBO + PREPOSICIÓN EN ELE: ANÁLISIS CONTRASTIVO DE LA INTERLENGUA EN APRENDIENTES DE NIVEL A2 Y B1

Grammatical collocations verb + preposition in Spanish as a foreign language: contrastive interlanguage analysis in learners of levels A2 and B1

RENÉ OPORTUS TORRES

Universidad de Concepción (Chile)
reneoportus@udec.cl

ANITA FERREIRA CABRERA

Universidad de Concepción (Chile)
aferreir@udec.cl

Resumen

Este estudio examina la frecuencia de colocaciones gramaticales verbo + preposición (ColGram v+p) en aprendientes anglófonos de Español como Lengua Extranjera (ELE) de nivel A2 y B1 bajo el modelo de Análisis Contrastivo de la Interlengua (ACI). Para ello se implementa una tipología colocacional sustentada en un criterio de fijación intermedia de estas unidades. Los resultados muestran mayor frecuencia en el nivel B1, pero no en la variedad de unidades, y un mayor número de usos correctos que en A2. Respecto de las diferencias encontradas según el patrón sintáctico-semántico, las variables competencia, complejidad del patrón y temáticas aparecen como factores importantes para explicarlas en el contexto del corpus de aprendientes examinado. Este trabajo aporta al estudio de las ColGram v+p, y de otros tipos (sustantivo/adjetivo + preposición) en ELE.

Palabras clave: Colocaciones gramaticales; verbo + preposición; ACI; ELE; escritura.

Abstract

This study examines the frequency of grammatical collocations verb + preposition (GramCol v+p) in L1-English learners of Spanish as a Foreign Language (SFL) of levels A2 and B1 under the model of Interlanguage Contrastive Analysis. To this end, a collocational typology based on a criterion of intermediate fixedness was implemented. The results show higher frequency in level B1, but not in the variety of the units, and a greater number of correct uses than in A2. Regarding found differences according to syntactic and semantic pattern, the variables of competence, pattern complexity and theme appear to play an important role in their explanation in the context of the learners' corpus examined. This work contributes to the study of GramCol v+p and other types (noun/adjective + preposition) in SFL.

Key words: Grammatical collocations; verb + preposition; ICA; SFL; writing skills.

INTRODUCCIÓN

La producción de la oración **me enamoré con el deporte*¹, en lugar de *me enamoré del deporte*, por parte de un aprendiente anglófono de ELE representa un error de tipo preposicional (*de* o *con*, en este caso), como se ha examinado en estudios de análisis de errores (AE) de aprendientes. Por otra parte, dicha desviación puede también analizarse desde una perspectiva fraseológica, en donde se consideraría que se trata de un error colocacional; es decir, el verbo *enamorarse* selecciona frecuentemente al colocativo *de* para expresar el significado de *sentir amor por*, lo que no se cumple en el ejemplo dado.

La combinación *enamorarse + de* se ajustaría plenamente a la definición de colocaciones gramaticales, en donde se les concibe como combinaciones semifijas y recurrentes constituidas por “una palabra dominante (sustantivo, adjetivo, verbo) y una preposición o una estructura gramatical como un infinitivo o cláusula” (Mi traducción) (Benson, Benson, & Ilse, 1986, ix). No obstante, a pesar de la relevancia de esta definición acerca de la lengua inglesa, esta no resultaría suficiente para delimitar lo que entendemos por colocaciones gramaticales en español (Rica, 2014). Debido a la importancia del estudio del léxico en la investigación de la adquisición de segundas lenguas (ASL), y en particular de las expresiones multipalabras, este artículo se focaliza en examinar las ColGram v+p en el ámbito de ELE desde una perspectiva fraseológica. Las ColGram², al igual que las colocaciones léxicas, se sitúan en un punto intermedio en un continuo de fijación entre las combinaciones libres y las locuciones idiomáticas (Henriksen 2013), y debido a su uso idiosincrático constituyen un área de dificultad para el aprendizaje de una L2 (Domínguez, 2014).

En ASL, la investigación pertinente las colocaciones se ha concentrado en el léxico de la lengua inglesa, y recientemente en ELE, con un foco en el desarrollo de aplicaciones informáticas de propósito lexicológico y pedagógico. En contraste, en las ColGram existe escasa investigación en ELE, a saber, Travalía (2006) y Domínguez (2014), lo que ameritaría su estudio. En esta senda, Rica (2010) destaca la importancia de las ColGram v+p, y por esta razón, la necesidad de estudiarlas, señalando que este tipo de patrones verbales constituye el grupo más amplio dentro de las ColGram.

Una razón que explicaría la limitada investigación de las ColGram en ELE se relaciona con la poca claridad en los rasgos que las diferencian de otros fenómenos lingüísticos. Se les ha igualado a casos ya explicados por la gramática tradicional, como es el sistema de régimen. En este sentido, la tipología propuesta por Travalía (2006) para el análisis de ColGram en español, basada en los criterios fraseológicos que caracterizan

¹ Ejemplo extraído del Corpus de Aprendientes CAELE, ver sección 2.3 para mayor información. El asterisco indica que la oración es agramatical, y que se presenta de la forma en que aparece en dicho corpus.

² Colocaciones gramaticales.

a las colocaciones léxicas, aparece como una gran contribución a su re-conceptualización, facilitando su identificación a un nivel operacional.

Este trabajo forma parte de una investigación mayor cuyo objetivo es contrastar el procesamiento cognitivo de las ColGram v+p en aprendientes de ELE. Este estudio corresponde a la primera etapa de dicha investigación, cuyo propósito es describir y comparar estas unidades, como un aspecto colectivo de la interlengua de estudiantes anglófonos, mediante el análisis de la escritura de aprendientes de niveles A2 y B1 en el marco de la Investigación de Corpus de Aprendientes y de una función diagnóstica del ACI (Sánchez, 2015). Se intenta responder a las interrogantes de cómo difiere: (1) la frecuencia de usos de ColGram v+p adecuados e inaceptables, (2) la variedad de unidades producidas, y (3) la frecuencia de los patrones sintáctico-semánticos vinculados a las ColGram v+p entre los niveles investigados. El estudio contempla la utilización de una tipología específica para clasificar las ColGram v+p según los patrones sintácticos y semánticos que caracterizan a estas unidades. Al respecto, Henriksen (2013, 36) señala la necesidad de estudios de mayor especificidad, indicando que “se han realizado pocos intentos por clasificar las colocaciones sistemáticamente en distintas subcategorías” (Mi traducción).

La primera parte del artículo reseña el marco teórico y el vacío en la investigación que justifica este trabajo. Luego se presenta el estudio mismo, con su marco metodológico, corpus, procedimientos, análisis, resultados y conclusiones.

1. MARCO TEÓRICO

1.1. EL ESTUDIO DE LAS COLGRAM V+P EN ELE

El interés por examinar las ColGram v+p deriva de los hallazgos de las primeras investigaciones de las unidades multipalabras en ASL. Estas se concentraron en estudiar las expresiones de mayor fijación bajo el supuesto que serían más útiles para el bagaje lingüístico de los aprendices. No obstante, estas mostraron que las unidades más frecuentes son en realidad las más transparentes y variables, lo que motivó un cambio de foco hacia las colocaciones y secuencias recurrentes que presentan mayor grado de sustitución y significación composicional. Así también, se reconoce la importancia de la competencia colocacional en la producción y comprensión de la L2.

Las colocaciones son estudiadas eminentemente desde 2 corrientes: un enfoque basado en frecuencia, y uno fraseológico. El primero tomaba como criterio básico la probabilidad estadística de que 2 o más palabras ocurriesen cercanas entre sí para determinar la ocurrencia de una colocación. Dicho criterio, de gran precisión, tiene la limitación de ignorar el aspecto semántico de estas unidades. El segundo enfoque ha delimitado las restricciones gramaticales y léxicas que distinguen a las colocaciones de otras unidades multipalabras. Estas se situarían en un punto intermedio entre las locuciones

idiomáticas y las combinaciones libres, mostrando la transparencia de estas últimas, y a su vez, los vínculos sintagmáticos de las locuciones idiomáticas. No obstante, este enfoque presenta la desventaja de que depende del trabajo manual para su identificación.

Desde la perspectiva fraseológica, la conceptualización de las ColGram en ELE ha sido influida por la distinción de Benson *et al.* (1986) entre colocaciones léxicas y gramaticales para el inglés, identificando ocho tipos de ColGram representadas por patrones sintácticos en donde una palabra de contenido léxico es seguida de una preposición o construcción gramatical. No obstante, se ha cuestionado que varias de estas categorías corresponden a fenómenos ya estudiados por la sintaxis.

Travalia (2006) plantea la necesidad de reconsiderar el concepto tradicional de ColGram, adhiriendo a la idea básica de que una colocación consiste en la ocurrencia frecuente, pero no obligatoria, de una palabra léxica y una preposición. En definitiva, propone una tipología para el español de 6 categorías definidas a base de criterios de alternancia entre preposiciones, la posible omisión de estas últimas, exigencia de una estructura particular después de la preposición y el efecto semántico de la ocurrencia preposicional versus construcciones alternantes sin preposición.

1.2 CORPUS ELECTRÓNICOS DE APRENDIENTES Y ANÁLISIS CONTRASTIVO DE LA INTERLENGUA

El objetivo principal de la investigación en ASL es develar los principios que rigen el aprendizaje de una L2. Siendo este un proceso mental, se debe estudiar de manera indirecta, por ejemplo, mediante el desempeño lingüístico. En esta senda, para investigar las ColGram v+p, debemos contar con evidencia lingüística que provea múltiples instancias de su uso. Los corpus electrónicos de aprendientes y de hablantes nativos constituyen recursos valiosos para acceder a las frecuencias de uso colocacional. Por medio de su análisis se pueden realizar contrastes para identificar el manejo y dificultades de los aprendientes. El modelo de ACI trata a la interlengua como una entidad autónoma e independiente, y se ha convertido en el procedimiento estándar para estos procedimientos. Su enfoque comparativo y naturaleza cuantitativa y cualitativa permite identificar los rasgos de distintas variedades de la interlengua; por ejemplo, entre diferentes orígenes lingüísticos, niveles de competencia, edad, etc., y así determinar la contribución de estos factores a un estadio en particular. También se pueden establecer comparaciones entre la interlengua y la variedad nativa para identificar similitudes y diferencias en el sobreuso o infrarrepresentación de algunos rasgos, indicadores de adquisición en etapas posteriores. Además de lo anterior, es posible combinar este enfoque con procedimientos del AE que entrega una metodología sistemática para la identificación, clasificación, descripción, y explicación de los errores (Ferreira & Elejalde, 2017). En definitiva, estos modelos combinados favorecen la descripción de la

interlengua de manera completa, más allá del error, señalando aciertos, y usos infrarepresentados o sobreutilizados.

1.3. LAS COLGRAM V+P EN ESTUDIOS PUBLICADOS

La investigación de las ColGram v+p en ELE ha sido escasa, lo que dificulta el establecimiento de datos de base para las comparaciones de este estudio. A saber, disponemos del estudio de Domínguez (2014), que compara la producción de colocaciones léxicas versus las ColGram en aprendientes de L1 valón y flamenco, y concluye que la frecuencia de ColGram supera las colocaciones léxicas; no obstante, es el segundo grupo el que muestra mayor cifra de colocaciones distintas. Este resultado coincide con Bahardoust y Moeini (2012) en inglés como L2. También en esa lengua como L2, se obtienen diferencias significativas en la producción y comprensión de ColGram (sin examinar las ColGram v+p) acorde la competencia (Alotaibi & Alotaibi, 2015), con mejor desempeño en el nivel avanzado. Con esto, constatamos que la investigación en ELE ha sido reducida, y que no se han analizado en profundidad las ColGram v+p en relación con la competencia.

2. EL ESTUDIO

Esta es una investigación descriptiva transversal, con un análisis mixto de datos. Estos se procesan cuantitativamente basados en la frecuencia, y cualitativamente durante la identificación, categorización, e interpretación de la aceptabilidad de las ColGram v+p. El objetivo general es comparar la frecuencia de usos adecuados e inaceptables, variedad de unidades producidas, y la frecuencia a lo largo de patrones descritos desde un punto de vista sintáctico-semántico entre los niveles de competencia estudiados.

2.1. PREGUNTAS DE INVESTIGACIÓN

1. ¿Difieren los niveles A2 y B1 en la frecuencia y variedad de ColGram v+p?
2. ¿Difieren los niveles A2 y B1 en la frecuencia de ColGram v+p usadas correctamente según el patrón sintáctico-semántico?

2.2. DESCRIPCIÓN DEL CORPUS DE APRENDIENTES Y LA MUESTRA

El Corpus especializado de Aprendientes de Español como Lengua Extranjera CAELE ha sido desarrollado en el marco de los proyectos de investigación FONDECYT N° 1140651 (Ferreira, 2014) y N° 1180974 (Ferreira, 2018). El Corpus CAELE (2017) cuenta con 729 textos redactados por aprendientes de diversos orígenes lingüísticos, los que han sido recopilados dentro de los proyectos señalados (Ferreira, 2014-2018).

Considerando los objetivos investigativos, se seleccionaron solo los textos de los 33 estudiantes de lengua materna inglés para conformar el subcorpus de aprendientes anglófonos. Su nivel de competencia (A2=19; B1=14) fue determinado con la Prueba de Multinivel con Fines Específicos Académicos (Ferreira 2016), instrumento alineado con los niveles del Marco Común Europeo de Referencia y administrado al comienzo del programa de enseñanza.

Las actividades de escritura se realizaron en el marco de la instrucción de la asignatura de español, directamente en el computador bajo la supervisión de un docente, sin acceso a diccionarios ni correctores. El tamaño total del subcorpus ascendió a 207 textos (116 en nivel A2, y 91 en B1). El total de palabras en el subcorpus ascendió a 44.624 palabras (A2=22.382, B1=22.242). La redacción de los textos estuvo dirigida por una función comunicativa, en estilo narrativo, descriptivo y argumentativo. Los temas abarcaron aspectos culturales y problemáticas actuales, con un foco en la apreciación de la realidad en los países de origen, con 20 temáticas en total, 16 de ellas temáticas comunes entre ambos grupos de competencia.

2.3. METODOLOGÍA DE LA INVESTIGACIÓN

La metodología del 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).

2.4. CONCEPTUALIZACIÓN Y TIPOLOGÍA DE COLGRAM V+P EN ESTE ESTUDIO

Nuestro concepto de ColGram recoge las características asociadas a las colocaciones léxicas. De esta forma, se definen operacionalmente como la combinación de un verbo pleno y una preposición que introduce un complemento preposicional no obligatorio. Estos complementos cumplirían la función sintáctica de un grupo argumental seleccionado semánticamente por el verbo, pero sintácticamente pueden omitirse al estar implícitos, o pueden alternar con otros grupos preposicionales o sintácticos para expresar un sentido similar.

Luego de la revisión preliminar del subcorpus y para describir las ColGram v+p, se desarrolló una tipología específica de análisis colocacional. Siguiendo a Travalia (2006), y las bases teóricas de la función del complemento preposicional (CPrep), se identificaron 4 patrones principales, con subtipos específicos definidos a base de los criterios de (1) 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 CPrep en algunos contextos.

A continuación, se especifican los patrones sintáctico-semánticos y subgrupos identificados en la tipología de ColGram v+p utilizada en este estudio:

- **Patrón 1.** Verbo + 2 ó más preposiciones alternantes con un significado similar.
 - Subtipo 1.1 Verbo + (CD³) + CPrep⁴ con preposiciones alternantes.
 - Subtipo 1.2 Verbo + CPrep con preposiciones alternantes.
 - Subtipo 1.3 Verbo pronominal + CPrep con preposiciones alternantes.
- **Patrón 2.** Verbo + 1 ó más de 1 preposición + estructura sintáctica específica.
 - Subtipo 2.1 Verbo + CPrep con estructura sintáctica específica (ej., infinitivo).
 - Subtipo 2.2 Verbo + CD + CPrep con estructura sintáctica específica.
 - Subtipo 2.3 Verbo + sintagma preposicional predicativo del sujeto o del CD
- **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.
 - Subtipo 3.2 Verbo no pronominal + CPrep que alterna con verbo no pronominal + CD, con diferencias semánticas importantes.
 - Subtipo 3.3 Verbo pronominal + CPrep que alterna con verbo no pronominal + CD, con significados cercanos
 - Subtipo 3.4 Verbo pronominal + CPrep que alterna con verbo no pronominal + CD, con diferencias semánticas importantes
- **Patrón 4.** Verbo + CPrep que puede aparecer tácito.
 - Subtipo 4.1 Verbo no pronominal + CD + CPrep que puede aparecer tácito.
 - Subtipo 4.2 Verbo no pronominal + CPrep que puede aparecer tácito.
 - Subtipo 4.3 Verbo pronominal + CPrep que puede aparecer tácito.

2.5. IDENTIFICACIÓN Y ACEPTABILIDAD DE LAS COLGRAM V+P

Las ColGram v+p fueron identificadas según los criterios de exigencia semántica y sintáctica dados por la definición operacional señalada antes. No obstante, debido a que estas secuencias se solapan con combinaciones v+p no colocacionales, identificamos de manera formal las concurrencias v+p excluidas del estudio: en la estructura interna en locuciones verbales y cuasilocuciones con verbo de apoyo; marcadores funcionales del CD, CF⁵ y complemento agente; CPrep obligatorio; perífrasis; complementos argumentales de ubicación complementos preposicionales adjuntos; oraciones copulativas; y secuencias verbo + locuciones adverbiales introducidas por preposición.

³ Complemento directo.

⁴ Complemento preposicional.

⁵ Complemento indirecto.

La clasificación de las ColGram v+p en adecuadas o inaceptables se basó en el criterio de combinatoria verbo + preposición esperada. Se consideraron adecuados los casos en que la preposición producida correspondía a la esperada (o esperadas) para expresar el sentido dado por el contexto del enunciado. Este proceso se basó en 2 fuentes: (1) la introspección del investigador, hablante nativo de español, respecto de la selección, interpretación y análisis de los datos a base de su conocimiento interiorizado de la L2 y fenómeno de estudio; (2) el Corpus del Español del Siglo XXI CORPES⁶, corpus de referencia representativo de la norma panhispánica del español actual (Sánchez, 2015).

Asimismo, adoptamos un enfoque fraseológico, focalizándonos solo en los errores que rompían el vínculo colocacional; ya fuese debido a la omisión o reemplazo de la preposición, o a la creación de una combinación v+p de naturaleza colocacional inaceptable 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 en el contexto de los enunciados donde ocurría la ColGram v+p, por ejemplo, a nivel de la palabra (errores ortográficos), o de la palabra y la oración (gramaticales o coherencia textual).

2.6. PROCESO DE ANOTACIÓN DEL SUBCORPUS

Siguiendo los procedimientos metodológicos de los Proyectos FONDECYT (Ferreira, 2014, 2018) en cuanto a las anotaciones taxonómicas e implementación en el *software* UAM (O'Donnell, 2008), se implementó una tipología específica para el estudio. Los textos se incorporaron en formato .txt con codificación UTF-8, y se etiquetaron manualmente acorde los atributos de aceptabilidad, tipo y subtipo sintáctico-semántico seleccionando los segmentos correspondientes. Asimismo, se ingresaron los datos de competencia de los sujetos para obtener los informes estadísticos descriptivos e inferenciales.

3. ANÁLISIS Y RESULTADOS

3.1. FRECUENCIA TOTAL Y VARIEDAD DE COLGRAM V+P SEGÚN LA COMPETENCIA

En total se identificaron 785 casos de ColGram v+p, que incluyen secuencias colocacionales inexistentes en español. De estos, 204 correspondieron a colocaciones distintas, así como *hablar de algo* y *hablar sobre algo*, o pares como *ir a - irse a*, diferenciadas por el incremento pronominal. En el nivel A2, el número de ColGram v+p fue de 363, con 122 colocaciones distintas; mientras que en B1 se encontraron 422 ocurrencias, con 140 colocaciones diferentes.

⁶ <http://web.frl.es/CORPES/>

Siguiendo a Orol & Alonso (2013), se calculó la frecuencia relativa de ColGram v+p respecto del total de palabras por nivel, obteniéndose 16,22 (A2) y 18,97 (B1) casos por mil palabras. Luego, para comparar la proporción de ColGram v+p versus el total de palabras en cada nivel, se aplicó la prueba Z (alfa=0.05), obteniéndose una diferencia entre ambas ($z=2.213$; $p=0,027$). Esto constata que el nivel B1 produjo más ColGram v+p que el nivel A2, en línea con Alotaibi & Alotaibi (2015) en inglés como L2 respecto de la diferencia según la competencia, con hallazgos de mayor frecuencia en los estudiantes de nivel más avanzado.

Asimismo, para determinar si existían diferencias según la competencia según las ColGram v+p distintas producidas, se calculó la proporción de ColGram v+p distintas versus el total producido en cada nivel, obteniéndose en A2=0,336, y en B1=0,332. Al comparar las proporciones se observó que no existían diferencias entre ellas ($z=0.128$; $p=0,899$). Esto sugiere que aunque la frecuencia tiende a aumentar en el nivel B1, esta no va necesariamente acompañada de un aumento en el repertorio que permita diferenciar la producción colocacional del nivel A2. Sobre ello, Domínguez (2014) sugiere que el rasgo gramatical de las ColGram (número fijo de preposiciones) establecería esta restricción. No obstante, esto derivaría de áreas temáticas subrepresentadas en el nivel B1, pues el grupo A2 redactó tareas de 20 temas distintos, en contraste con los 16 temas del nivel B1.

También se examinaron las 10 mayores frecuencias para determinar si algunas ColGram v+p específicas concentraban parte del total producido; el uso repetido puede dar indicio del manejo según nivel (Tabla 1). La colocación con mayor frecuencia en ambos niveles es *ir a* (A2=48 oc.⁷; B1=26 oc.) ($z=0.027$; $p=0,001$). También aparecen en ambos niveles colocaciones como *tratarse de*, *ayudar a algo* y *hablar sobre*, no obstante, con frecuencias porcentuales y posiciones distintas en el ordenamiento grupal.

Otro dato interesante surge al comparar las frecuencias acumuladas en la quinta y décima posición del ordenamiento descendente de frecuencias. En la quinta posición las frecuencias acumuladas alcanzan una proporción relativamente cercana en ambos grupos (A2=21,8%; B1=22,0%), y en la décima posición sucede algo similar (A2=28,7%; B1=30,6%). Esto nos lleva a plantear que la producción de ambos grupos no parecería estar caracterizada por una sobrerrepresentación de unas pocas colocaciones en particular, sino que estas tienden a distribuirse de manera proporcional en forma paralela en ambos subcorpus; no obstante, aparecen excepciones (*ir a*), con una frecuencia considerablemente mayor en el nivel A2 que en el nivel B1.

⁷ Ocurrencias.

Tabla 1: Frecuencia y porcentaje acumulado de colocaciones más usadas según nivel

Posición	ColGram v+p	A2 (N 363)			B1 (N 422)			
		Frecuencia absoluta	%	% acumulado	ColGram v+p	Frecuencia absoluta	%	% acumulado
1	ir a	48	13,2	13,2	ir a	26	6,2	6,2
2	viajar a	9	2,5	15,7	tratar de (asunto)	22	5,2	11,4
3	tratarse de (asunto)	8	2,2	17,9	hablar sobre	17	4,0	15,4
4	ayudar a (+infinitivo)	8	2,2	20,1	escribir sobre	16	3,8	19,2
5	compartir con	6	1,7	21,8	ayudar a (+infinitivo)	12	2,8	22,0
6	hablar sobre	5	1,4	23,1	tratarse de (asunto)	8	1,9	23,9
7	apoyar a (+infinitivo)	5	1,4	24,5	hacer de (materia)	8	1,9	25,8
8	quedarse en	5	1,4	25,9	hablar de	8	1,9	27,7
9	cantar sobre	5	1,4	27,3	saber sobre	6	1,4	29,1
10	escribir sobre	5	1,4	28,7	interesarse en	6	1,4	30,6

3.2. FRECUENCIA DE COLGRAM V+P SEGÚN COMPETENCIA Y PATRÓN SINTÁCTICO-SEMÁNTICOS

La prueba Chi Cuadrado obtenida de UAM⁸ se empleó para responder a la interrogante de cómo difieren los 2 niveles en los usos adecuados e inaceptables, y según subgrupo sintáctico-semántico en el caso de los usos adecuados. En primer lugar, se identificó el número de usos adecuados e inaceptables del total de 785 ColGram v+p. En el análisis global, el total de usos adecuados correspondió a 490 (62,4%), superando los usos inaceptables con una frecuencia de 295 (37,6%) ($z=11.695$; $p=0.000$); lo que coincide con lo reportado por Bahardoust y Moeini (2012) en inglés como L2.

Al comparar los usos adecuados e inaceptables según competencia, se observaron diferencias entre ambos niveles (Tabla 2). Los usos adecuados del nivel B1 superan al nivel A2, y los usos inaceptables del nivel A2 superan al nivel B1. Esto se enmarca dentro de lo esperado en términos de que un mejoramiento en el nivel lingüístico debería ir asociado a mayor precisión, y reducción en la tasa de error.

Tabla 2: Frecuencia y porcentaje de usos adecuados e inaceptables según competencia

	A2 (N=363)		B1 (N=422)		Chisqu	Signif
	Frecuencia	%	Frecuencia	%		
Usos adecuados	210	57,85	280	66,35	6,01	+++
Usos inaceptables	153	42,15	142	33,65	6,01	+++

Focalizándonos en los usos adecuados para identificar las tendencias acorde los 4 patrones sintáctico-semántico, al comparar las frecuencias obtuvimos que la mayor

⁸ Significancia estadística es expresada mediante: (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.

proporción se presentaba en el patrón 1 (353 oc., 72,0%), superando al 4 (91 oc., 18,6%), 3 (26 oc., 5,3%), y 2 (20 oc., 4,1%) ($z=1.681$; $p=0.000$; $z=2.145$; $p=0.000$; $z=2.190$; $p=0.000$, respectivamente), en forma descendente. Por otra parte, la proporción del patrón 4 sobrepasa a los patrones 2 ($z=7.156$; $p=0.000$) y 3 ($z=6.404$; $p=0.000$), siendo estos dos últimos los que no muestran diferencia entre las proporciones. Luego, al desglosar los datos anteriores acorde la competencia, la frecuencia del nivel B1 presentaba una diferencia estadística levemente significativa respecto del nivel A2 en los patrones 1 y 2, superando al nivel A2 (Tabla 3). En cuanto a los patrones 3 y 4, no se observaron diferencias estadísticas significativas según nivel.

Tabla 3: Frecuencia y porcentaje de usos adecuados por nivel de competencia

Patrón sintáctico semántico	A2 (N=363)		B1(N=422)		Chisqu	Signif.
	Frecuencia	%	Frecuencia	%		
Patrón 1	150	41,32	203	48,10	3,63	+
Patrón 2	5	1,38	15	3,55	3,73	+
Patrón 3	13	3,58	13	3,08	0,15	
Patrón 4	42	11,57	49	11,61	0,00	
Total usos adecuados	210		280			

En el análisis por subtipo sintáctico-semántico (Tabla 4), sobresale la frecuencia de usos aceptables del subtipo 1.1 con la mayor frecuencia (39,0%), y descendiendo, los subtipos 1.2 (27,5%), 4.1 (9,4%), y 1.3 (5,5%). El resto de subtipos muestra frecuencias inferiores al 5,5%. Al comparar los subtipos sintáctico-semánticos según la competencia, se obtienen diferencias estadísticas significativas en los subtipos 1.1, 1.2, y 2.2. De estos casos, la frecuencia del nivel B1 supera al nivel A2 en los subtipos 1.1 y 2.2, mientras que en el subtipo 1.2 es el grupo A2 el que muestra mayor frecuencia. Los demás subgrupos no muestran diferencias entre los niveles.

Tabla 4: Frecuencia y porcentaje de usos adecuados por tipo sintáctico-semántico

Patrón sintáctico semántico		A2 (N=363)		B1 (N=422)				Total (N=490)	
Patrón	Subtipo	Frec.	%	Frec.	%	Chisqu	Signif.	Frec.	%
1	1.1	57	15,7	134	31,8	27.31	+++	191	39,0
	1.2	83	22,9	52	12,3	15.23	+++	135	27,6
	1.3	10	0,3	17	2,6	7.04		27	5,5
Sub-total		150		203				353	
2	2.1	2	0,6	3	0,7	0.08		5	1,0
	2.2	1	0,3	11	2,6	7.04	+++	12	2,4
	2.3	2	0,6	1	0,2	0.51		3	0,6
Sub-total		5		15				20	
3	3.1	4	1,1	4	1,0	0.05		8	1,6
	3.2	5	1,4	4	1,0	0.32		9	1,8
	3.3	1	0,3	4	1,0	1.39		5	1,0
	3.4	3	0,8	1	0,2	1.34		4	0,8
Sub-total		13		13				26	
4	3.1	21	5,8	25	5,9	0.01		46	9,4
	3.2	9	2,5	17	4,0	1.46		26	5,3
	3.3	12	3,3	7	1,7	2.24		19	3,9
Sub-total		42		49				91	
Total		210		280				490	

Un aspecto que puede dar cuenta de las diferencias anteriores es el uso recurrente de algunas colocaciones específicas. La Tabla 5 muestra la frecuencia de colocaciones con 3 o más ocurrencias del subtipo 1.1 y 1.2 en los niveles A2 y B1 por nivel. En el subtipo 1.1, la mayor frecuencia del grupo B1 puede explicarse por la producción de colocaciones específicas, con usos sobrerrepresentados: *escribir sobre* (16 oc.), *hablar sobre* (17 oc.), y *tratar de* (22 oc.), superando a las del nivel A2 (5, 5, y 3 oc., respectivamente). En el subtipo 1.2, tenemos una situación similar, pero inversa considerando la competencia. El grupo A2 supera al nivel B1 debido a la frecuencia de 2 colocaciones en particular: *ir a* y *viajar a*. Estas, en conjunto, explican 57 ocurrencias del total de 83 en el nivel A2, superando ampliamente al nivel B1 (26, y 4, respectivamente).

Tabla 5: Frecuencia de ColGram v+p más frecuentes subtipos 1.1 y 1.2 según competencia

Patrón 1. Subtipo 1.1			Patrón 1. Subtipo 1.2		
ColGram v+p	Frecuencia		ColGram v+p	Frecuencia	
	A2 (N=57)	B1 (N=134)		A2 (N=83)	B1 (N=52)
aprender de (asunto)	3	4	ir a (destino)	48	26
aprender sobre (asunto)	5	4	regresar a (destino)	3	-
cantar de (asunto)	5	3	venir a (destino)	-	4
conocer de (asunto)	-	3	venir de (procedencia)	4	-
enseñar sobre (asunto)	-	3	viajar a (destino)	9	4
enviar a (destino)	-	3	volver a (destino)	4	-
escribir sobre (asunto)	5	16			
escuchar en (medio)	-	4			
hablar de (asunto)	-	8			
hablar sobre (asunto)	5	17			
hacer de (materia)	-	6			
saber sobre (asunto)	3	6			
tratar de (asunto)	3	22			
tratar sobre (asunto)	-	5			

Finalmente, el subtipo 2.2 incluye un grupo potencial reducido de ColGram: *alentar a*, *animar a*, *apoyar a*, *impulsar a*, *ayudar a*, *presionar a*, etc. La diferencia de frecuencia entre el nivel A2 (1) y el nivel B1 (11) podría originarse en la complejidad que estas colocaciones presentarían para el nivel A2, la que se expresa en la posición indirecta del CPrep respecto del verbo, y que exige un infinitivo después de la preposición.

5. CONCLUSIÓN

Este artículo abordó el tema de las ColGram v+p en ELE con el objetivo de describir y comparar estas unidades en la producción escrita en estudiantes anglófonos de niveles A2 y B1. Para ello se conceptualizó e implementó una tipología respecto de los patrones sintácticos y semánticos que describen este tipo específico de colocaciones siguiendo el criterio de fijación intermedia de estas combinaciones.

Entre los hallazgos destacan: (1) que el nivel B1 produce un mayor número de unidades que el nivel A2; no obstante, la proporción de ColGram v+p distintas no difiere entre ellos, (2) la mayor frecuencia de colocaciones específicas en el nivel B1 que en el nivel A2 (*escribir sobre*, *hablar sobre*, *tratar de*), (3) la similitud entre los 2 niveles respecto de la colocación *ir a* con la mayor frecuencia en ambos grupos, y (4) la representación similar de las 5 y 10 mayores frecuencias en ambos grupos, que explican alrededor del 22,0% y 30,0%, respectivamente, de las frecuencias totales.

El nivel B1 presenta también un mayor número de colocaciones correctas que el nivel A2, lo que es esperable respecto de la competencia. En cuanto a los usos correctos según

patrón sintáctico-semántico, las mayores frecuencias ocurren en los patrones 1 y 4. Además, se observa que el nivel B1 tiene una mayor frecuencia en los patrones 1 y 2. Respecto de los subtipos sintáctico-semánticos, el nivel B1 supera al nivel A2 en los subtipos 1.1 y 2.2, este último se explicaría por la mayor complejidad de este patrón, lo que presentaría una mayor dificultad para el nivel A2. Por otra parte, el nivel A2 obtiene mayor número en el subtipo 1.2, explicado por la alta representación de la colocación *ir a*. En definitiva, las diferencias se originarían en los factores de competencia, complejidad de los patrones sintáctico-semánticos y temáticas vinculadas a las tareas de escritura desarrolladas.

Este estudio no da cuenta exhaustiva de la naturaleza y etiología de los errores en la producción de ColGram v+p, ni del aspecto de naturalidad del uso de estas unidades respecto de la producción nativa, resultados a presentarse en la etapa siguiente de la investigación mayor en la que se sitúa este estudio. Por otra parte, la concepción y tipología colocacional utilizada permitió examinar las ColGram v+p, un fenómeno cubierto escasamente por la investigación en ELE.

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” del autor principal, dirigida por 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.

OBRAS CITADAS

- Alotaibi, Addullah; Alotaibi, Maye (2015). “The Comprehension and production of English grammatical collocations by Kuwaiti EFL learners”, *International Journal of English Language Teaching* 13(3), 26-39.
- Bahardoust, Maryam; Moeini, Mahammad (2012). “Lexical and grammatical collocations in writing production of EFL learners”, *The Journal of Applied Linguistics* 5(1), 61-86.
- Benson, Morton, Benson, Evelyn. & Ilson, Robert (1986). *The BBI Combinatory Dictionary of English. A guide to word combinations*. Amsterdam: John Benjamins.
- Domínguez, Meritxell. U. (2015). Las colocaciones en un corpus de aprendices valón y flamenco. En *La enseñanza de ELE centrada en el alumno* (pp. 977-986). Asociación para la Enseñanza del Español como Lengua Extranjera.

- Ferreira, Anita (2014). *Proyecto de investigación FONDECYT - CONICYT N°1140651: El feedback correctivo escrito directo e indirecto en la adquisición y aprendizaje del español como lengua extranjera*. Universidad de Concepción, Chile 2014-2017.
- Ferreira, Anita (2014-2018). *Corpus de Aprendices de Español Lengua Extranjera, CAELE*, Universidad de Concepción, Chile
- Ferreira, Anita (2016). *Prueba de Multinivel con Fines Específicos Académicos, Programa de Español como Lengua Extranjera ELE-UdeC*, Universidad de Concepción, Chile.
- Ferreira, Anita (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, Anita; Elejalde, Jessica (2017). “Análisis de errores recurrentes en un Corpus de aprendices de español como lengua extranjera (Corpus CAELE)”. *Revista Brasileira de Lingüística Aplicada* 17(3), 509-538.
- Granger, Sylviane (2015). “Contrastive interlanguage analysis: A reappraisal”, *International Journal of Learner Corpus Research* 1(1), 7-24.
- Henriksen, Birgid (2013). Research on L2 learners’ collocational competence and development – a progress report. En C.Bardel, C.Lindquist & B. Laufer (Eds.). *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis* (pp 29-57). Amsterdam: Eurosal.
- O'Donnell, Michael (2008). “The UAM CorpusTool: Software for corpus annotation and exploration”. *Proceedings of the XXVI Congreso de AESLA*, Almería, Spain, 3-5 April
- Orol, Ana; Alonso, Margarita (2013). “A comparative study of collocations in a native corpus and a learner corpus of Spanish”. *Procedia* 95, 563-570.
- Rica, Juan (2010). “Colocaciones gramaticales en la producción escrita de estudiantes universitarios españoles”. *Reduca (Filología)* 2(1), 1-25.
- Sánchez, Ana (2015). “Las funciones diagnóstica y evaluativa del análisis contrastivo de la interlengua del español basado en corpus”. *Linred*, 13.
- Travalia, Carolina (2006). “Las colocaciones gramaticales en español”. *Anuario de estudios filológicos*. 29, 279-295.