

Análisis de Errores en el Corpus CAELE: estudio sobre la concordancia gramatical en el verbo "ser" en aprendientes francófonos y anglosajones

Error Analysis in the CAELE Corpus: study on the grammatical agreement in the verb "ser" in Francophone and Anglo-Saxon learners

Anita Ferreira

Universidad de Concepción
aferreir@udec.cl

Jessica Elejalde

Universidad de Concepción
jelejalde@udec.cl

Lorena Blanco

Universidad Católica de la Santísima Concepción
lblanco@ucsc.cl

RESUMEN

En este artículo se presenta un estudio sobre Análisis de Errores Asistido por el Computador. Para ello, se considera la identificación de los errores y la tendencia de uso erróneo en Español como Lengua Extranjera (ELE) en el nivel de competencia A2 en dos subcorpus del corpus CAELE, un subcorpus de 798 enunciados textuales de aprendientes anglosajones de ELE y otro subcorpus de 625 de aprendientes francófonos. Los objetivos generales son determinar las frecuencias de errores de concordancia gramatical en el verbo ser y comparar las semejanzas y diferencias observadas entre aprendientes francófonos y anglosajones. Los resultados muestran una mayor dificultad para los aprendientes anglosajones en la selección de las categorías gramaticales de género y número en los adjetivos, especialmente en la concordancia de los atributos. En cambio, los aprendientes del francés presentan una frecuencia de errores notoriamente menor en dichos aspectos gramaticales. Los errores podrían atribuirse a la transferencia directa de reglas gramaticales de la L1, incidiendo en la precisión gramatical de los enunciados con el verbo ser. Estos hallazgos contribuyen a dar mayor solidez a las tendencias de la literatura especializada.

Palabras clave: Análisis de Errores Asistido por el Computador, Corpus de Aprendientes, concordancia gramatical, precisión gramatical, Español como Lengua Extranjera.

ABSTRACT

This article presents a study on Computer Aided Error Analysis. For this, the identification of errors and the tendency of misuse in Spanish as a Foreign Language (ELE) at the A2 proficiency level is considered in two subcorpus of the CAELE corpus, a subcorpus of 798 textual sentences of English language learners of ELE and another subcorpus of 625 French-speaking learners. The general objectives are to determine the frequencies of grammatical concordance errors in the verb ser and to compare the similarities and differences observed between French-speaking and Anglo-Saxon learners. The results show greater difficulty for Anglo-Saxon learners in the selection of the grammatical categories of gender and number in the adjectives, especially in the concordance of the

attributes. In contrast, French learners show a noticeably lower frequency of errors in these grammatical aspects. The errors could be attributed to the direct transfer of grammar rules from L1, affecting the grammatical precision of the sentences with the verb to be. These findings contribute to giving more solidity to the trends in the specialized literature.

Keywords: Computer-Aided Error Analysis, Computer Learner Corpora, grammatical agreement, grammatical precision, Spanish as Foreign Language.

1. INTRODUCCIÓN

En la literatura especializada de adquisición de segundas lenguas (ASL), los estudios sobre la temática de los errores causados por transferencia lingüística negativa se han conducido a través de dos enfoques metodológicos, el análisis de errores asistido por el computador (del inglés *Computer Aided Error, CEA*) y el análisis contrastivo de la interlengua (ACI) (Gilquin, 2008; Granger, 1996). Granger (2009) sostiene que el ACI es una técnica heurística que permite descubrir rasgos de la lengua de los aprendientes y luego analizarlos desde una perspectiva de la L2. Los planteamientos actuales van más allá del error, señalando los aciertos, frecuencia de uso, usos infrautilizados, y la sobreutilización de formas de la L2 (Ishikawa, 2013; Sánchez, 2015).

En dicho contexto, el presente estudio se enmarca en las áreas de Corpus de Aprendientes en formato computacional (del inglés, *Computer Learner Corpora, CLC*). Los corpus computacionales fueron originados en la década de 1990, tras el interés por crear y generar recursos que fueran de utilidad para los aprendientes de lenguas. Como resultado, la consolidación de la Lingüística de Corpus y la evolución de la tecnología y de las metodologías de investigación permitieron el uso de dichos corpus como herramientas de investigación, posibilitando el desarrollo de nuevos tipos de estudios tales como los análisis frecuencias de coaparición o de patrones de uso lingüístico por parte de grupos de aprendientes (Granger, 2004, 2012, 2015; Hasselgard y Johansson, 2011; Sánchez, 2015).

Los corpus de aprendientes se construyen a partir de textos auténticos producidos por personas que están aprendiendo una determinada lengua con diferentes lenguas maternas (L1) y distintos grados de conocimiento de la lengua meta. Los textos se recolectan de acuerdo a criterios específicos cumpliendo con condiciones de tipología textual, ámbito geográfico, temporalidad, etc., que permiten que el corpus se constituya en una muestra representativa de una lengua particular o de un subconjunto de la misma (Bautista, 2019). Estos corpus permiten que el investigador detecte las fortalezas y debilidades del grupo lingüístico estudiado, por lo tanto, los resultados se consideran una fuente valiosa de datos lingüísticos (Buyse, 2011).

Las investigaciones basadas en corpus de aprendientes (Campillos, 2014; Granger, 2015 y Lozano, 2015) han contribuido en la descripción de la interlengua relacionada con los aspectos lingüísticos que se aprenden antes que otros, sobre cuáles son más difíciles de aprender y las diferencias con los hablantes nativos (HN), por medio de análisis contrastivos. Los corpus de aprendientes son especialmente útiles a la hora de investigar el proceso de aprendizaje y constituyen una fuente empírica de datos para líneas de

investigación como el CEA (del inglés *Computer aided error analysis*) y los estudios en análisis de interlengua (Cruz Piñol, 2012). Al mismo tiempo, estas descripciones mejoradas de la lengua del aprendiente pueden ser utilizadas en la enseñanza de las L2 (Granger, 2012; Pastor Cesteros, 2004).

Para encauzar la investigación en este ámbito interdisciplinario, hemos estado implementando un corpus de aprendientes de Español como Lengua Extranjera, el cual hemos denominado, CAELE. Este corpus se constituye en una colección abierta que año tras año considera la recolección y el procesamiento de nuevos textos. El conjunto de textos recolectados ha sido anotado y procesado a través del software *Sketch Engine* (Kilgarriff et al., 2004; 2014), configurándose en una colección de 1217 textos producidos en un ambiente presencial en un aula virtual. Los aprendientes, que escribieron dichos textos, provienen de diferentes países, tales como, Francia, Bélgica, Alemania, Sudáfrica, Estados Unidos, Inglaterra, Brasil, Portugal, entre otros. Los estudiantes pertenecen a diferentes áreas de estudio: ingeniería, matemáticas e industria, administración y astrofísica, áreas más relacionadas con las ciencias exactas y el uso del lenguaje matemático. También provienen de las áreas de las humanidades y ciencias sociales.

El CAELE permite observar distintas problemáticas lingüísticas y realizar análisis léxico-gramatical, que puede incluir listas de frecuencias, colocaciones, expresiones multipalabras, palabras clave, n-gramas, entre otros. La problemática que se aborda en el presente artículo se relaciona con las dificultades que se observan en las estructuras nominales y adjetivas con el verbo *ser* en aprendientes de ELE de nivel de competencia A2 con L1 francés e inglés. En general, los estudios relacionados con el verbo *ser* abordan el tema sobre el carácter intercambiable de los verbos copulativos *ser* y *estar* (Baralo, 1998; Guijarro-Fuentes, 2008). La distinción entre el uso del verbo *ser* y *estar* es uno de los temas gramaticales de mayor dificultad para los aprendientes de ELE, no solo para aprendientes francófonos o anglosajones, sino también para estudiantes de muchas otras lenguas. Tanto en el francés como en el inglés sólo existe un verbo *être* y *to be*, respectivamente, para lo que en español tenemos el verbo *ser* y el verbo *estar*.

Otra problemática interesante de explorar y describir es la relativa a la concordancia gramatical en las estructuras de sujeto y predicativas con el verbo *ser*. En efecto, una de las áreas que presenta bastante dificultad para los aprendientes de ELE anglosajones lo constituyen los errores de concordancia (Fleissig, 2009; Méndez y Rojas, 2013; Naranjo, 2009; Paquet, 2017; Prieto, 2011). Como bien lo señala Pastor Cesteros (2001) se debería analizar los problemas de transferencia en la concordancia gramatical, especialmente en los niveles de competencia principiante, para evitar la fosilización de dichos errores. La L1 y el nivel de competencia en la L2 se consideran factores influyentes en la transferencia. En consecuencia, determinar en qué medida la transferencia y el nivel de competencia influyen en los errores de concordancia gramatical es todavía un tema no resuelto que requiere de mayor investigación (Paquet, 2017).

En el presente artículo se muestran los resultados de un estudio descriptivo de carácter mixto, cuantitativo y cualitativo para delimitar los errores más frecuentes cometidos por aprendientes de ELE francófonos y anglosajones de nivel A2 en las estructuras de sujeto y atributivas con el verbo *ser*, así como las principales semejanzas y diferencias observadas entre ambos grupos. El artículo se organiza en las siguientes secciones: En la sección 2, nos referimos a los fundamentos teóricos en materia de la concordancia gramatical y los estudios sobre aprendientes de ELE. En la sección 3, abordamos el estudio sobre un análisis de errores de concordancia gramatical. En la sección

4, presentamos los resultados cuantitativos y cualitativos. Finalmente, presentamos algunas conclusiones sobre los avances y logros obtenidos en esta investigación.

2. FUNDAMENTOS TEÓRICOS

2.1. La concordancia gramatical en español

El Diccionario de la Real Academia Española (2014: 531) define la concordancia gramatical como la "congruencia formal que se establece entre las informaciones flexivas de dos o más palabras relacionadas sintácticamente". A su vez, en el Diccionario Panhispánico de Dudas (2005: 158), se explica la concordancia como "la coincidencia obligada de determinados accidentes gramaticales (género, número y persona) entre distintos elementos variables de la oración". En dicho diccionario se distinguen dos tipos de concordancias:

a) nominal, en la cual el sustantivo establece concordancia de género y número con sus determinantes, tales como artículos y adjetivos. Por ejemplo, "Enriqueta es malintencionada". En esta oración el número y el género del sujeto se corresponden con el atributo nominal.

b) verbal, establecida entre el sujeto y el verbo de una oración. Por ejemplo, "Los alumnos están muy bien". La frase nominal desempeña función sintáctica de sujeto y concuerda en número plural con el verbo.

La problemática de la concordancia se rige por algunas reglas generales. Entre ellas:

1) La coordinación de dos o más sustantivos o pronombres en singular, siempre que cada uno de ellos se refiera a un ente distinto, forma un grupo que concuerda en plural con el adjetivo o el pronombre, o con el verbo del que son sujeto. Por ejemplo, "Los estudios y los viajes son compatibles".

2. La coordinación de dos o más sustantivos o pronombres de diferente género gramatical forma un grupo que concuerda en masculino con el adjetivo o con el pronombre: "En la universidad, las flores y los murales eran coloridos".

La gramática española reconoce el valor de dos propiedades de esta naturaleza que están implicadas en la concordancia. Estas son género y número.

2.1.1. Género gramatical en español

Según el Manual de la Nueva Gramática de la Lengua Española (2010: 23), el género es "una propiedad gramatical de los sustantivos y de algunos nombres que incide en la concordancia con los determinantes, los cuantificadores y los adjetivos o los participios". Montes se refiere a este fenómeno como:

El género es un sistema de clasificación de los nombres en grupos o clases que pueden incluir o no la clase 'sexo'. Además, agrega que la concordancia es la que hace posible dicha clasificación. Cuando se hace alusión al género, el nombre es el que lleva la información gramatical, pero cuando actúa dentro de una expresión debe concordar con las otras categorías gramaticales tales como el adjetivo o el pronombre que lo acompañan (Montes, 1997: 144-145).

2.1.2. Número gramatical en español

El Diccionario de la Real Academia (2014: 1444) define en una de sus acepciones el concepto de número como la "categoría gramatical que en el nombre y el pronombre

expresa la referencia a una o a varias entidades, y que en otras clases de palabras se manifiesta a través de la concordancia". En el Manual de la Nueva Gramática de la Lengua Española este concepto significa:

una propiedad gramatical característica de los sustantivos, los nombres, los adjetivos, los determinantes (en el sentido amplio, que abarca los cuantificadores) y los verbos. Se presenta en dos formas: singular (árbol, quien, grande, este, come) y plural (árboles, quienes, grandes, estos, comen) (RAE, 2010: 35).

Sobre la base de las conceptualizaciones expuestas, se puede señalar que ambas propiedades gramaticales tienen un sistema de valoración dual, ya que distinguen entre el singular y el plural para el caso de número, y entre el masculino y el femenino en el caso del género. Ambas categorías se manifiestan por medio de desinencias o sufijos flexivos que se agregan a la raíz de los sustantivos, adjetivos, entre otros (por ejemplo, o/a-o/e-o/a en chico/chica, cliente/clienta, león/leona para el género).

2.2. Estudios sobre la concordancia gramatical en aprendientes de Español como lengua Extranjera

En los estudios sobre la concordancia basados en el análisis contrastivo en la adquisición de la sintaxis del español pueden destacarse las investigaciones de la evolución de la interlengua en hablantes de lengua nativa inglesa (Liceras, 1983), estudios sobre el análisis de errores en los tiempos de pasado (Fernández, 1997); estudios de la interlengua y análisis de errores de hablantes de L1 alemana (Rakasaeder y Schmidhofer, 2014; Vázquez, 1991), serbocroata (Santos Gargallo, 1993) o de diversas L1 (Fernández, 1997). Asimismo, existen estudios de la interlengua del español en aspectos morfológicos como la formación de palabras (Baralo, 1996) o en otros como los clíticos, las oraciones relativas especificativas o el parámetro del sujeto nulo (Liceras, 1996); y, finalmente, trabajos de lingüística contrastiva del español con lenguas como el portugués, el inglés o el chino (Penadés, 1999). Aun cuando, han sido pocos los estudios que se focalizan en el fenómeno de la concordancia, no obstante, dichos estudios contribuyen en los análisis de la conformación de la interlengua y los procesos que subyacen a este proceso.

En el estudio de Vázquez (1991), se explica que en la producción de la concordancia en el sintagma nominal, el aprendiente activa diversas operaciones mentales. Estas "operaciones cognitivas, por un lado, a partir del conocimiento de las reglas morfológicas de formación de género y número, y operaciones psicológicas, por otro, en la medida en que se presupone un análisis semántico del sustantivo, especialmente en el caso de los colectivos" (Pastor Cesteros, 2001: 10). En dicha explicación Vázquez (1991) considera que este tipo de operaciones ocurren dada la experiencia y el conocimiento que tiene el aprendiente de su L1, lo que podría resultar en apariciones de interferencias cuando no haya equivalencias entre ambas lenguas (Pastor Cesteros, 2001). En este contexto, Vázquez (1991) señala que en la mayoría de los casos los errores de concordancia se pueden explicar en términos de uso de estrategias de simplificación por parte del aprendiente y en su investigación, se refiere a la interferencia de la L1, además de otros factores mencionados.

Por su parte, Fernández (1997) analiza la interlengua de aprendientes de diversas L1, profundiza diversos errores gramaticales cometidos, señalando los problemas de concordancia con mayor frecuencia. El estudio se realizó sobre la base de cuatro grupos de

L1 alemana, japonesa, árabe y francesa, de distintos niveles de adquisición, a continuación se referencia lo indicado por Fernández (1997): 1) frecuencia de errores de concordancia cuando el sustantivo y el elemento que concuerda con este no se sitúan contiguamente, 2) el cambio de género del elemento que no concuerda es más el masculino por el femenino que viceversa, 3) respecto de la concordancia en número, hay preferencia por la forma del singular frente a la de plural; la falta de concordancia se produce mayormente en la categoría de adjetivo; y un factor que la favorece es precisamente la no distinción entre adverbio y adjetivo. Según Fernández (1997), el aumento de errores, en un determinado momento, no es signo de regresión sino de progreso, pues ello implica en el aprendiente un proceso más creativo, de ensayar y practicar atreviéndose a arriesgar más en la nueva lengua.

Con otro enfoque, Pastor Cesteros (2001) lleva a cabo un análisis de un corpus multilingüe y de niveles iniciales, intermedios y avanzados compuesto por pruebas de español como LE. Los resultados en relación con la concordancia gramatical establecen cuatro tipos de errores que explican el avance del nivel de competencia y la conformación de la interlengua. Estos corresponden a: 1) la concordancia entre determinante y sustantivo, 2) sujeto y atributo, 3) adjetivo y sustantivo. Estos hallazgos determinan cómo en el nivel inicial se cometen errores prácticamente en todos los ítems, mientras que en los siguientes niveles va aumentando el dominio de este aspecto, hasta el punto de que en el nivel superior el error aparece sólo en un tercio de los ítems y en un porcentaje mucho menor. En lo referido a la concordancia entre sustantivo y adjetivo (sean estos elementos de un único SN o sujeto y atributo, respectivamente), se especifica que a partir del nivel avanzado desaparecen gradualmente los problemas con adjetivos terminados en -e (y de un modo definitivo ya en el nivel superior). Se concluye que en la concordancia sustantivo/adjetivo (tanto en el seno del SN, como en la relación sujeto/atributo), surgen errores en el nivel inicial, los cuales permanecen en el nivel intermedio en los casos de adjetivos acabados en -e, cuando hay alternancia en la adscripción del determinante y cuando hay inversión de las normas generales respecto al género y se encuentra ante sustantivos masculinos acabados en -a que se hacen, por tanto, concordar en femenino. Este tipo de problemas disminuye considerablemente en el nivel avanzado y prácticamente es inexistente en el superior.

Por su parte Aránzazu (2004) especifica la importancia de estudiar la concordancia de género de los sustantivos inanimados de género no semántico, como *problema*, *mano*, etc., con sus modificadores en aprendientes de ELE anglosajones. La autora señala que la adquisición de la concordancia de los sustantivos problemáticos resulta complicada para cualquier alumno de ELE, cuya L1 carezca de esta categoría morfosintáctica. Sobre la base de dichos resultados Aránzazu (2004) propone la enseñanza de estos aspectos gramaticales centrada en el procesamiento del input y su realce para facilitar el aprendizaje en el ELE.

En estudios más recientes en otras lenguas, como en el caso del análisis de la concordancia en la expresión escrita de aprendientes serbios de ELE (Rubinjoni, 2014), se señala que las causas de los errores podrían residir en la transferencia del comportamiento gramatical de las unidades de una lengua a las de la otra, lo que supone considerar el conocimiento de la lengua y su influencia en la conformación de oraciones y discurso con sentido.

En las investigaciones de Elejalde y Ferreira (2016) y Ferreira y Elejalde (2020b) referidos al análisis de frecuencia de errores de transferencia de la L1 al ELE de aprendientes anglófonos y germanoparlantes las autoras concluyen que la tendencia de la falsa selección de la concordancia gramatical en el atributo de oraciones copulativas y la elección del género

gramatical se corrobora con los hallazgos en el estudio de Pastor Cesteros (2001) y Rakasaeder y Schmidhofer (2014). No obstante, los errores del género se observaron con mayor frecuencia en el alemán, dadas las características del género morfológico en dicha lengua. En esta última se transfieren los rasgos de género propios del alemán, induciendo al error de falsa selección de género gramatical y morfológico tanto en el sintagma nominal como verbal.

Finalmente, en el estudio de Agudelo (2018) enfocado en identificar y describir los errores y causas del verbo *ser* y *estar* en estudiantes francófonos de ELE, se parte de la base de que el uso entre estos dos verbos del español es una dificultad para los estudiantes que poseen un sistema verbal de atributo diferente. En sus hallazgos la investigadora señala con respecto al verbo *ser* en el nivel A2 que, la mayoría de los errores cometidos por este grupo de aprendientes está presente en las estructuras conformadas por verbo + atributo adjetivo (que solo van con *ser*). Afirma también que en este nivel de competencia del español la concordancia en la conjugación constituye una dificultad lingüística y no la selección entre el verbo *ser* y *estar*. En relación con las causas, se atribuye la confusión morfológica, definida como los errores de conjugación en los cuales no hay concordancia entre la forma y el sujeto. Por ejemplo, "*¿Dónde está ellas?*"

3. EL ESTUDIO

El diseño del estudio que se presenta a continuación es de carácter descriptivo con un enfoque de análisis de datos mixto: cuantitativo y cualitativo. En cuanto a los datos de corte cuantitativo, permitirán identificar las frecuencias y sistematicidad de los errores, así como también determinar las tendencias de uso de acuerdo con las funciones gramaticales o sintácticas que correspondan. Con respecto a los datos cualitativos, estos proporcionarán las explicaciones sobre el posible origen de los errores según los análisis de la influencia directa de la L1. Se busca avanzar en responder la pregunta de investigación sobre ¿Cuáles son las principales dificultades que manifiestan los alumnos universitarios francófonos y anglosajones de ELE en tareas de escritura académica?

El enfoque metodológico de la investigación se enmarca en los procedimientos de la lingüística de corpus, el análisis de errores y el análisis contrastivo de interlengua. Los procedimientos se basan en los lineamientos propuestos en el Proyecto Fondecyt No.1180974. A continuación, presentamos los objetivos generales y específicos del estudio:

Objetivos generales

1. Determinar las frecuencias de errores de concordancia gramatical de género y número en estructuras oracionales con el verbo *ser* a nivel de sujeto, sujeto verbo y atributos nominales y adjetivales, producidas por aprendientes francófonos y anglosajones de nivel A2 de ELE.
2. Comparar las principales diferencias y semejanzas observadas en los errores de aprendientes francófonos y anglosajones en los contextos de estudio con el verbo *ser*.

Objetivos específicos

- 1.1 Determinar y explicar las mayores frecuencias de errores en la concordancia gramatical en estructuras oracionales con el verbo *ser* en aprendientes francófonos.
- 1.2 Determinar y explicar las mayores frecuencias de errores en la concordancia gramatical en estructuras oracionales con el verbo *ser* en aprendientes anglosajones.

3.1. El Corpus de Aprendientes de Español como Lengua Extranjera, CAELE

El CAELE es una colección de 1217 textos producidos en un ambiente presencial a través de un aula virtual, procesados a través del software *Sketch Engine*. Estos textos han sido recolectados durante los periodos de clases entre los años 2014 y 2019 y se distribuyen en los niveles de competencia del ELE, A2 y B1. CAELE está conformado por 265.362 palabras, por 296.136 tokens y por 14.902 oraciones. La recolección de los textos se realizó durante los cursos de ELE ofrecidos semestralmente cada año en una universidad chilena. La extensión es de 150 a 200 palabras (A2) y de 250 a 300 (B1). Las temáticas de escritura se determinaron mediante un análisis de necesidades y consideraron temas del ámbito de español para fines académicos científicos y humanísticos: literatura, ciencia y tecnología, cultura latinoamericana, etc. Las secuencias textuales recolectadas en los textos responden a la estructura narrativa, descriptiva y argumentativa (Ferreira y Elejalde, 2019). Para ratificar los niveles de competencia que los aprendientes declaran en una encuesta de registro, se aplicó la prueba de multinivel de competencia (Ferreira, 2016). Los aprendientes que escribieron los textos son 210 sujetos y provienen de los programas de intercambio de una universidad chilena en el marco de la movilidad estudiantil.

Para este estudio se crearon 2 subcorpus de enunciados textuales a partir del CAELE, relacionados con el objeto de estudio de esta investigación. Un subcorpus que está constituido por 625 enunciados textuales con el verbo *ser* producidos por aprendientes francófonos y otro subcorpus está constituido por 798 enunciados textuales producidos por aprendientes anglosajones.

A continuación se explican los procedimientos de extracción de los subcorpus a partir del corpus CAELE y los procesos de anotación a partir de la taxonomía de errores y procesamiento de los datos de dichos corpus.

3.2. Implementación de los subcorpus

El corpus CAELE fue implementado en el software *Sketch Engine*. Para extraer los 2 subcorpus francófono y anglosajón se ingresó la búsqueda de concordancias con el verbo *ser* en la herramienta *Word Sketch*. En seguida, se seleccionaron los atributos más representativos (nominales y adjetivos) para finalmente filtrar los resultados por la L1 de los aprendientes (francés e inglés) y nivel de competencia A2. En la Figura 1 se muestra parte del resultado de la búsqueda en la interfaz concordancia, correspondiente al atributo nominal. Se ilustra la interfaz con la que funciona el análisis del software *Sketch Engine*. Los enunciados pueden analizarse de acuerdo con su concordancia, los tipos de palabras, el contexto de la oración y el número de frecuencias o apariciones en las consultas por cada caso.



Figura 1. Concordancias con el verbo *ser* en *Sketch Engine*

Por otra parte, en el Gráfico 1 se muestra la conformación de los subcorpus del CAELE para el análisis de los errores de concordancia gramatical. Cada subcorpus de enunciados se subdivide en dos tipos de atributo, el nominal y el adjetivo. Respecto de las oraciones con atributo nominal el total corresponde a 663 frecuencias, de las cuales 349 corresponden a la producción de hablantes nativos de inglés y 314 de hablantes nativos de francés. En relación con las oraciones con atributo adjetivo se refleja un total de 760 frecuencias distribuidas en 449 frecuencias de hablantes nativos de inglés y 311 de francés. En total se consideraron para este análisis 1423 oraciones con concordancia gramatical de atributo con el verbo copulativo *ser*.

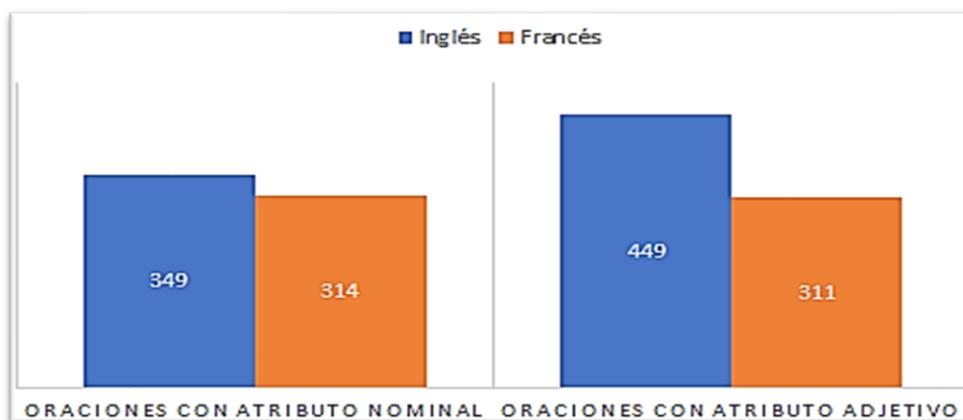


Gráfico 1. Distribución por lengua materna

En el Gráfico 1 se puede observar la distribución de las oraciones con atributo nominal y con atributo adjetivo en las lenguas en estudio. Se observa que en inglés hay mayor uso de oraciones con atributo adjetivo (449 apariciones) respecto de las oraciones con atributo nominal con 349 usos, mientras que en francés se usa equilibradamente ambos tipos de oraciones con un total de 314 de tipo nominal y 311 de tipo atributo adjetivo.

3.3. La anotación

Respecto del análisis y clasificación de los errores se consideró la taxonomía propuesta para el etiquetado de errores del criterio etiológico (Ferreira y Elejalde, 2020a). Para la identificación de los errores de concordancia gramatical se incluyó 9 tipos de etiquetas que se especifican en la Tabla 1. Las etiquetas se forman por dos bases principales, la primera (FS) nomenclatura que responde a *falsa selección*, en términos de la modificación de la lengua objeto de estudio y la segunda (CG) nomenclatura que significa *concordancia gramatical* para referirse a la temática central de este estudio. Las otras nomenclaturas dependen del tipo de concordancia en estudio. Las etiquetas corresponden a: 1) FS-CG-GEN, falsa selección de la concordancia gramatical de género en el sujeto, 2) FS-CG-NUM, falsa selección de la concordancia gramatical de número, 3) FS-CG-SUJ-VSER, falsa selección de la concordancia gramatical entre el sujeto y el verbo ser, 4) FS-CG-ATRIB-GEN-FA falsa selección de la concordancia gramatical del atributo en el género frase adjetival, 5) FS-CG-ATRIB-NUM-FA falsa selección de la concordancia gramatical del atributo en el número frase adjetival, 6) FS-CG-ATRIB-GEN-FN falsa selección de la concordancia gramatical del atributo en el género frase nominal, 7) FS-CG-ATRIB-NUM-FN falsa selección de la concordancia gramatical del atributo en el número frase nominal, 8) FS-CG-GEN falsa selección de la concordancia gramatical en el género en el atributo interno, 9) FS-CG-NUM falsa selección de la concordancia gramatical en el número en el atributo interno.

Tipos CG	Etiqueta	Descripción
1) Sujeto	FS-CG-GEN	Errores de concordancia gramatical de género y de número en la frase nominal que funciona como sujeto de la oración con verbo <i>ser</i>
	FS-CG-NUM	
2) Sujeto-predicado	FS-CG-SUJ-VSER	Error gramatical de concordancia gramatical de número entre el sujeto y el verbo <i>ser</i>
3) Atributo CG	FS-CG-ATRIB-GEN-FA	Errores en el atributo frase nominal o en el atributo frase adjetivo que generan falta de concordancia gramatical con el verbo <i>ser</i> o con el sujeto del enunciado.
	FS-CG-ATRIB-NUM-FA	
	FS-CG-ATRIB-GEN-FN	
	FS-CG-ATRIB-NUM-FN	
4) Atributo interno	FS-CG-GEN	Errores de concordancia gramatical de género y número a nivel interno en el atributo ya sea frase adjetival o nominal.
	FS-CG-NUM	

Tabla 1. Especificación de la taxonomía y las etiquetas para identificación y clasificación de errores (Ferreira y Elejalde, 2020b)

Como se observa en la Tabla 1, de acuerdo con la problemática de este estudio se delimitaron cuatro tipos de concordancia gramatical para el análisis, las que corresponden a los siguientes ejemplos:

(1) Tipo 1. La concordancia entre los elementos del sintagma nominal en función de sujeto de la oración copulativa. Un ejemplo de este caso corresponde a: "*Hay muchas tipos de música que existe en Latino América, algunos ejemplos **son** la salsa, la bachata y la cumbia*" (S43A2T4I2016)¹ y "*Me gusta la comida del país. Los paisaje **son** muy captivo y magnifico*" (S66A2PRUMF2018).

En los enunciados puede observarse la falta de concordancia en el sujeto, tanto a nivel de género "*algunos ejemplos*" como en el número "*los paisaje*".

(2) Tipo 2. La concordancia gramatical de número entre el sujeto y el verbo *ser*. Los ejemplos para este caso son: "*En el centro hay muchas departamentos de gobierno y muchas edificios que es importante de historia de Chile*" (S51A2POSDIFI2016).

En el enunciado se muestra la falta de concordancia gramatical entre el sujeto pronombre relativo (que) el cual se refiere a "muchas edificios" y la forma verbal del verbo *ser* "es".

(3) Tipo 3. La concordancia gramatical de género y número del atributo respecto del sujeto. Los ejemplos son: "*En mi país la música folclórica **es** sencillo y en general hay solo un cantante*" (S11A2POSINMI2014) y "*Todas las personas **son** diferente por eso hay personas a que gustan la música folclórica*" (S19A2PREF2015).

En los enunciados se ilustra la falta de concordancia gramatical entre el sujeto y el atributo del verbo *ser*. Es decir, hay una falsa selección del género entre "la música" y "sencillo". Además, se observa falta de concordancia de número entre "las personas" y "diferente".

(4) Tipo 4. La concordancia gramatical del atributo considerando el sintagma nominal o adjetival respecto de su función como atributo. Por ejemplo: "*Las ferias con soporte local **son** buenas lugares para la gente que quieren vivir con las ventas de artesanía*" (S11A2T1I2014)

En este último ejemplo, se muestra la falta de concordancia gramatical de género entre el sujeto y el atributo de adjetivo en "las ferias" y "buenas lugares".

Con la delimitación de estos cuatro tipos de concordancia gramatical se busca determinar no sólo la concordancia relacionada con el sujeto y el atributo, sino también las concordancias internas a nivel del sujeto y sujeto verbo *ser*. De esta forma, se cautela la revisión de la concordancia que podría afectar el sentido parcial o completo de la oración.

3.4. Procesamiento de los datos del subcorpus CAELE con el verbo *ser*

Para realizar el análisis de errores, como ya se señaló, se extrajo a través del software *Sketch Engine* un subcorpus de concordancias textuales conformado por 780 concordancias de aprendientes anglófonos y 625 concordancias de francófonos. Una vez obtenido los

subcorpus, se procedió a la identificación y clasificación de los errores. Para ello, se llevó a cabo de forma manual un registro de los errores que se encontraban y las frecuencias. Este registro se alojó en Excel para el posterior tratamiento de los datos.

Con los registros de las frecuencias en el Excel se calculó el total de errores y las frecuencias relativas por cada caso; al mismo tiempo, se determinaron los errores más frecuentes y cuáles de estos tenían mayor incidencia en la conformación de las oraciones copulativas del verbo *ser*. Respecto a la distribución de los enunciados textuales para el análisis, se consideró la variable *investigador*, por lo que estos fueron distribuidos y asignados de manera aleatoria para ser analizados por el equipo de investigación.

4. RESULTADOS Y DISCUSIÓN

De acuerdo con nuestro Objetivo General 1 se determinó los errores más frecuentes para establecer las mayores dificultades en el uso de la concordancia gramatical observadas en los enunciados de los aprendientes. Para ello, se realizó el conteo de los errores identificados en los cuatro tipos especificados anteriormente. Para responder a los objetivos específicos 1 y 2 se distribuyeron los resultados generales acorde con la variable L1 francés e inglés.

4.1. Errores más frecuentes en la concordancia gramatical con el verbo *ser*

El total de errores identificados fue de 282 errores en ambas lenguas en estudio, francófonos y anglosajones. En la Tabla 2 se presenta el total de los errores en los cuatro tipos de concordancias analizadas. En el Tipo 1, sujeto, se encontraron 55 errores en la concordancia gramatical de género y número en el sintagma nominal del sujeto. Seguido de este se observa el Tipo 2 (sujeto-predicado) con un total de 35 errores de concordancia entre el sujeto y el verbo *ser*. En el Tipo 3, se hallaron 136 errores en la concordancia gramatical entre el atributo y el sujeto de la oración copulativa. Por último, en el Tipo 4 hay un total de 55 errores en la concordancia gramatical de género y número del atributo a nivel del sintagma nominal o adjetival. De esta manera, puede observarse con precisión qué ocurre tanto a nivel del sintagma como la estructura completa de la oración copulativa.

Por otra parte, el análisis de estas frecuencias arroja como error más frecuente, con 80 errores, la falsa selección de concordancia gramatical de número en el atributo en la frase adjetival (*FS-CG-ATRIB-NUM-FA*), seguido de la falsa selección de concordancia gramatical de género en el atributo interno (*FS-CG-GEN*) con 48 errores y la falsa selección de concordancia gramatical de género presente en el sujeto de la oración (*FS-CG-GEN*) con 42 errores. Estas frecuencias reflejan cuáles son las problemáticas más relevantes en el CAELE en cuanto a la concordancia gramatical en oraciones copulativas con el verbo *ser*.

No. Tipo	Tipos CG	Etiqueta	Total	Fr. %
1	Sujeto	FS-CG-GEN	42	15%
		FS-CG-NUM	13	5%
2	Suj-pred	FS-CG-SUJ-VSER	35	12%
3	Atributo	FS-CG-ATRIB-GEN-FA	28	10%

	Concordancia Gramatical en FAdj y FN	FS-CG-TRIB-GEN-FN	12	4%
		FS-CG-TRIB-GEN-FA	3	3%
		FS-CG-TRIB-GEN-FN	2	2%
4	Atributo interno	FS-CG-GEN	17	18%
		FS-CG-NUM	3	3%
Total de todos los tipos			282	100%

Tabla 2. Total de frecuencias de errores según los 4 tipos de concordancias gramaticales en estudio

4.2. Frecuencia de errores de concordancia gramatical en aprendientes francófonos

De acuerdo con el Objetivo Específico 1, en esta subsección se presentan los errores más frecuentes de los aprendientes de L1 francés y su distribución según los tipos de concordancia en estudio (ver Tabla 3).

El error más frecuente corresponde a la falsa selección en la concordancia gramatical de número en la frase adjetival con un total de 30 apariciones (32%) en el Tipo 3 (atributo CG). Seguido se encuentran dos errores con una frecuencia menor, la falsa selección de la concordancia gramatical entre el sujeto y el verbo copulativo con un total de 18 errores (19%) en el Tipo 2 y la falsa concordancia en el género dentro del sujeto con 12 errores (13%) en el Tipo 1.

De acuerdo con lo señalado por Doquin de Saint Preux y Sáez (2014) la concordancia de la categoría gramatical de número es uno de los errores más frecuentes cometidos por aprendientes francófonos de nivel de competencia avanzada. En esta categoría gramatical los errores se podrían considerar errores de transferencia dado que los francófonos comenten de manera recurrente dicho tipo de error en su propio idioma. Según Doquin de Saint Preux y Sáez (2014) la concordancia de número, en francés, presenta el mismo mecanismo que en español, sin embargo, al no pronunciarse la "s" final del sustantivo tienden a omitirla, también de manera equivocada, en la lengua escrita.

No.Tipo	Tipos CG	Etiqueta	L1 Francés	Fr. %
1	Sujeto	FS-CG-GEN	12	13%
		FS-CG-NUM	6	6%
2	Suj-pred	FS-CG-SUJ-VSER	18	19%
3	Atributo CG	FS-CG-TRIB-GEN-FA	3	3%
		FS-CG-TRIB-GEN-FN	3	3%
		FS-CG-TRIB-GEN-FA	30	32%
		FS-CG-TRIB-GEN-FN	2	2%
4	Atributo interno	FS-CG-GEN	17	18%
		FS-CG-NUM	3	3%
Total de errores			94	100%

Tabla 3. Frecuencia de errores en la concordancia gramatical de oraciones copulativas con el verbo *ser* en aprendientes francófonos

Otro error frecuente es la falsa selección del género gramatical en la composición interna del atributo de la frase nominal con un total de 17 errores (18%) en el Tipo 4. Este tipo de errores en la categoría de género gramatical se pueden explicar por la transferencia desde la L1, ya que los aprendientes le adjudican al sustantivo el género que tiene dicha palabra en francés (Doquin de Saint Preux y Sáez, 2014).

A continuación, se presentan ejemplos y la explicación de los casos más frecuentes en los aprendientes francófonos:

- 1) Ejemplo Tipo 1, FS-CG-GEN: "El segundo lugar que tienes absolutamente que ver es la ciudad de Valparaiso. *Este ciudad es* el puerto el más importante del Chile y se puede encontrar muy grande barcos" (S52A2PRUMF2017).

En este ejemplo no hay concordancia de género gramatical a nivel de la frase nominal del sujeto, probablemente debido al desconocimiento del uso del adjetivo determinante "este".

- 2) Ejemplo Tipo 2, FS-CG-SUJ-VSER: "*Mi rutina en Chile son* muy diferentes de mi vida" (S90A2PRU2F2019).

En este ejemplo hay una falsa selección del número gramatical del verbo, el cual no concuerda con el sujeto que está en singular. Según afirma Agudelo (2018) estos errores entre la concordancia del verbo se podrían explicar en términos de confusión morfológica donde el aprendiente no identifica claramente la categoría de número en el verbo y su concordancia con el atributo.

- 3) Ejemplo Tipo 3, FS-CG-TRIB-NUM-FA "*Los autos electricos o que funcionan con agua son* muy caro" (S58A2PRU2F2017).

En el ejemplo 3, se observa la falsa selección de la concordancia gramatical del número en la frase adjetiva del atributo. Dicho atributo no concuerda con el sustantivo "*los autos*". Este error podría deberse a la distancia entre la frase adjetiva "muy caro" y la frase nominal referenciada por el pronombre relativo "que". El aprendiente se confunde por la subordinación. Estas confusiones igualmente ocurren en el español cuando hay una distancia mayor entre el verbo *ser* y los predicativos. También se observa este fenómeno en oraciones con pronombres relativos y referenciales en su función anafórica.

- 4) Ejemplo Tipo 4, FS-CG-GEN "*La pintura describe el mesclar de los racas en sur america . es* un pintura de acrylico" (S19A2T3F2015).

Finalmente, en el ejemplo 4 hay una falsa selección del género en la función interna del atributo "un pintura de acrylico". Este error puede deberse al desconocimiento del género gramatical del sustantivo induciendo al error. El desconocimiento del género de los sustantivos representa uno de los aspectos más problemáticos en el aprendizaje del español como LE (Pastor Cesteros, 2001).

4.3. Frecuencia de errores de concordancia gramatical en aprendientes anglosajones

De acuerdo con el Objetivo Específico 2, en esta subsección se presentan los errores más frecuentes de los aprendientes de L1 inglés y su distribución acorde con los tipos de concordancia en estudio (ver Tabla 4). Los errores más frecuentes identificados de mayor a menor en la Tabla 5 se distribuyen de la siguiente forma: (1) del Tipo 3, falsa selección de concordancia gramatical de número en la frase adjetiva del atributo con un total de 50 errores (27% del total de los errores); (2) del Tipo 4, la falsa selección de género en la conformación interna del atributo con un total de 31 errores (16%); (3) del Tipo 1, la falsa selección del género en la concordancia interna del sujeto con un total de 30 errores (16%); (4) del Tipo 3, la falsa selección de concordancia gramatical de género con un total de 25 errores (13%) en la frase adjetiva del atributo, y (5) del Tipo 3, la falsa selección de la concordancia gramatical de número en la frase nominal del atributo de la oración copulativa con un total de 15 errores (8%).

No. Tipo	Tipos CG	Etiqueta	L1 Inglés	Fr. %
1	1.Sujeto	FS-CG-GEN	30	16%
		FS-CG-NUM	7	4%
2	2.Suj-pred	FS-CG-SUJ-VSER	17	9%
3	3.Atributo CG	FS-CG-ATRIB-GEN-FA	25	13%
		FS-CG-ATRIB-NUM-FA	50	27%
		FS-CG-ATRIB-GEN-FN	9	5%
		FS-CG-ATRIB-NUM-FN	15	8%
4	4.Atributo interno	FS-CG-GEN	31	16%
		FS-CG-NUM	4	2%
Total de errores			188	100%

Tabla 4. Frecuencia de errores en la concordancia gramatical de oraciones copulativas con el verbo *ser*, aprendientes anglosajones

A continuación, se presentan algunos ejemplos con su respectiva explicación:

- 1) Ejemplo del Tipo 1, FS-CG-GEN: "*La ambiente es bueno y personas necesita reservado una mesa*" (S51A2T1I2016).

Se observa una falsa selección del género a nivel del sujeto, "la ambiente". Este tipo de error se puede explicar también (al igual que en el caso de los aprendientes francófonos) por un desconocimiento del género del sustantivo. Errores muy frecuentes según la literatura especializada.

- 2) Ejemplo Tipo 3, FS-CG-ATRIB-NUM-FA: "*Los otros estudiantes de intercambio son muy amable y nosotros trabajamos junto para hacer muchas cosas*" (S76A2PRUMI2018).

En este ejemplo se evidencia la falta de concordancia en el número gramatical entre el sujeto y el atributo, "*los otros estudiantes son muy amable*". En inglés los adjetivos no tienen rasgos morfológicos de género y número por lo que se podría interpretar que el

aprendiente, en este caso, transfiere esta regla gramatical al español induciendo al error en la concordancia.

- 3) Ejemplo Tipo 3, FS-CG-ATRIB-GEN-FA: "El Gran Cañon **es** muy bonita y gente de todos lugares viene para verlo" (S40A2PREI2016)

En este ejemplo hay una falsa selección del género gramatical en el atributo induciendo al error en la concordancia con el sujeto, "El Gran Cañón es muy bonita". Esto podría deberse al desconocimiento del género del sustantivo o bien a las estrategias intralingüísticas que operan en la interlengua del aprendiente. Es decir, muchas de las producciones de los aprendientes de lenguas se enfocan en la reducción o invención, la generalización, en la evasión o en la superproducción. En este caso, podría explicarse en términos de reducción de la morfología del adjetivo, induciendo al error en la falta de concordancia gramatical.

- 4) Ejemplo Tipo 3, FS-CG-ATRIB-GEN-FN: "*La artesanía es una producto que es*" (S25A2T1I2015)

Finalmente, en el ejemplo 4 se presenta un error en la concordancia de género gramatical del atributo. Al parecer, el sujeto intenta concordar el atributo nominal "una producto" con el sujeto "la artesanía". Este tipo de error podría explicarse en términos de la estrategia interlingüística de generalización de las reglas que resultan de fácil comprensión para el aprendiente.

En síntesis: a diferencia de los aprendientes francófonos, los anglosajones tienden a transferir las estructuras lingüísticas de su L1 dada la distancia entre ambas lenguas. La explicación en los errores de los francófonos podría estar relacionada con la transferencia del sistema lingüístico que se asemeja o se acerca a la conformación del español. Estas dos fuentes del error muestran la interferencia del conocimiento del sistema lingüístico de la L1 como estrategia de aprendizaje para sortear con éxito los vacíos de aprendizaje o conocimiento lingüístico a nivel del género gramatical de los sustantivos. Los resultados aquí encontrados guardan relación con estudios de transferencia, reflejándose en su mayoría tendencias similares en el uso incorrecto de la concordancia gramatical (Agudelo, 2018; Doquin de Saint Preux y Sáez, 2014; Elejalde y Ferreira, 2016; Ferreira y Elejalde, 2017; Pasquet, 2017; Pastor Cesteros, 2001).

4.3. Comparación de errores de concordancia entre aprendientes francófonos y anglosajones

En cuanto al Objetivo General 2, esto es, "comparar las principales diferencias y semejanzas observadas en los errores de aprendientes francófonos y anglosajones en los contextos de estudio con el verbo SER", a continuación se presenta los resultados de las frecuencias por cada L1, considerando la relación entre la frecuencia relativa (respecto del número de errores por lengua) y la frecuencia aumentada (referida a la proporción entre los errores de todo el corpus y los errores encontrados en cada lengua). Para ello, también se ha intentado identificar las posibles fuentes que inducen al error respecto de la conformación del sistema lingüístico tanto del inglés como del francés. El vasto número de adjetivos que existen en español, el aspecto semántico de los mismos, la pragmática y la

variedad geográfica son factores que hacen complejo el campo atributivo y contribuyen a la dificultad en el uso del verbo *ser* (Agudelo, 2018).

Tipos CG	Etiqueta	L1	Fr. %	Fr.	L1 Inglés	Fr. %	Fr.
		Francés		Aum			Aum
1.Sujeto	FS-CG-GEN	12	13%	4%	30	16%	11%
	FS-CG-NUM	6	6%	2%	7	4%	2%
2.Suj-pred	FS-CG-SUJ-VSER	18	19%	6%	17	9%	6%
3.Atributo CG	FS-CG-ATRIB-GEN-FA	3	3%	1%	25	13%	9%
	FS-CG-ATRIB-NUM-FA	30	32%	11%	50	27%	18%
	FS-CG-ATRIB-GEN-FN	3	3%	1%	9	5%	3%
	FS-CG-ATRIB-NUM-FN	2	2%	1%	15	8%	5%
4.Atributo interno	FS-CG-GEN	17	18%	6%	31	16%	11%
	FS-CG-NUM	3	3%	1%	4	2%	1%
Total de errores		94	100%	33%	188	100%	67%

Tabla 5. Frecuencia de errores por L1 en estudio (Francés e inglés)

En la Tabla 5 se puede ver que en los aprendientes de L1 francés hubo un total de 94 errores que corresponde al 33% en relación con el 100% del total de los errores. Por el contrario, la producción de errores de los aprendientes de L1 inglés arrojó un total de 188 errores lo que representa el 67% en relación con el 100% del total de los errores. Este primer análisis refleja que hay una notoria diferencia en cuanto a los porcentajes de errores, evidenciando que se presentan mayores problemáticas de concordancia en la producción de aprendientes anglosajones.

En cuanto al porcentaje de errores en el francés, puede inferirse que las similitudes entre esta lengua y el español favorecen al aprendizaje de este último.

Además, puede observarse que en cada lengua los errores más frecuentes corresponden a:

- 1) En L1 inglés se identifican mayores frecuencias en 4 tipos de errores: (1) errores de concordancia de falsa selección en el atributo en el número en la frase adjetiva con 50/188 errores (Tipo 1). (2) Errores de falsa selección del género en la frase interna del atributo con 31/188 errores (Tipo 4). (3) Errores de concordancia de género gramatical a nivel interno de la FN del sujeto con 30/188 errores (Tipo 1) y por último (4) Errores en la concordancia de género gramatical entre sujeto y el atributo con 25/188 errores (Tipo 3).
- 2) En la L1 francés, puede constatarse tres errores más frecuentes: (1) la falsa selección de número en la concordancia del atributo y el Sujeto con 30/94 errores (Tipo 3). (2) La falsa selección de número del sujeto y predicado del verbo *ser* con un total de 18/94 errores (Tipo 2). Y por último, (3) el error referido a la concordancia de género gramatical a nivel interno del atributo con 17/94 errores (Tipo 4).

Con estos resultados obtenidos, se puede inferir que los errores más frecuentes corresponden al desconocimiento del género gramatical del sustantivo en español, lo que induce al error de concordancia. Asimismo, podría explicarse que en el inglés los errores son más frecuentes en cuanto al género dado que esta lengua no hace una distinción entre el género y el número de los sustantivos, lo que supondría una dificultad mayor para los aprendientes anglosajones (Elejalde y Ferreira, 2016). A diferencia en el francés, algunos de los errores de género y número se reflejan porque muchos de estos no coinciden con el género ni el número en francés y en español, lo que podría determinarse como una interferencia directa.

Los hallazgos aquí encontrados pueden ser corroborados con otros estudios respecto de la complejidad que tienen los rasgos morfológicos y sintácticos del género y número en el español y la distancia de esta lengua con lenguas que no pertenecen al grupo romance (Agudelo, 2018; Elejalde y Ferreira, 2016; Paquet, 2017). Por su parte, Bond et al. (2011) examinaron la influencia de la L1 a partir de un grupo de estudiantes de L1 inglés de nivel competencia inicial y delimitaron que tres tipos de concordancias: de número sujeto verbo, de número sustantivo adjetivo y de género sustantivo adjetivo. Los resultados mostraron mayor sensibilidad de los aprendientes para percibir errores de concordancia de número que de género.

Para una mejor visualización de los resultados logrados con respecto al Objetivo General 2 y poder observar las principales diferencias y semejanzas en los errores en ELE según las L1 en estudio se presentan los Gráficos 1 y 2, que reflejan las tendencias de uso erróneo en ambos grupos de aprendientes.

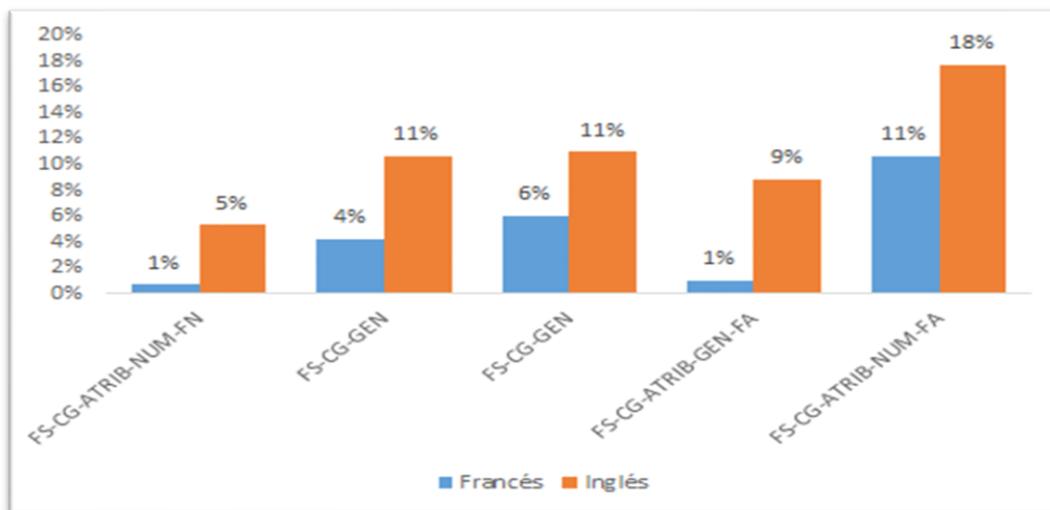


Gráfico 3. Diferencias en las frecuencias de errores entre los francófonos y anglosajones

En el Gráfico 3, se constata las principales diferencias entre los 5 errores más frecuentes del análisis de la concordancia gramatical en los enunciados textuales con el verbo *ser*. Según la distribución de mayor a menor de los porcentajes de errores se visualiza (1) la falsa selección del número gramatical en la concordancia de atributo en la frase adjetiva con un 18% en L1 inglés y 11% en L1 francés, reflejando 7 puntos de diferencia. (2) El error de falsa selección de la concordancia de género gramatical en la frase adjetiva

con un 1% de L1 francés y un 9% de la L1 inglés, con una diferencia de 8 puntos. (3) El error de falsa selección de concordancia gramatical de género a nivel interno en el atributo con un 6% en L1 francés y un 11% en L1 inglés con una diferencia de 5 puntos. (4) La falsa selección del género en la concordancia gramatical a nivel interno del sujeto con un 4% en L1 francés y con un 11% en L1 inglés, marcando una diferencia de 7 puntos. Finalmente, (5) la falsa selección de la concordancia gramatical de número en el atributo de la frase nominal con un 1% en el francés y un 5% en el inglés (4 puntos de diferencia).

Como explicación de estas diferencias, se puede sostener que para los aprendientes anglosajones hay mayores dificultades en la selección del género gramatical y de número en los adjetivos, especialmente en la concordancia de los atributos, dado que, en inglés, como ya se ha dicho, la estructura de los sintagmas nominales no concuerda ni en género ni en número con el sustantivo. Es decir, que estos errores podrían atribuirse a la transferencia directa de estas reglas gramaticales, afectando la precisión gramatical de la oración copulativa del verbo *ser*. Respecto del francés, no se observa que estas problemáticas sean muy frecuentes en las oraciones, dado que esta lengua tiene rasgos morfológicos de género y número que coinciden con las reglas del español (Elejalde y Ferreira, 2016; Agudelo, 2018). No obstante, existe otra posible explicación en la cual se le atribuye la distancia lineal o sintáctica que expone que, a mayor distancia entre los elementos de una oración, la concordancia podría presentar dificultades (Paquet, 2017). Si se remite a los ejemplos anteriores, muchas de las oraciones tienen varios modificadores del sujeto que podrían afectar la concordancia tanto a nivel de sintagma nominal o adjetival como el sujeto y el atributo.

En cuanto a las similitudes en los resultados encontrados entre las frecuencias de errores en ELE, el Gráfico 4 muestra las principales estructuras que fueron menos problemáticas para ambos grupos de aprendientes. Esto es: 1) la falsa selección del número gramatical en la conformación interna del atributo con un porcentaje de un 1%. 2) La falsa selección de la concordancia de número gramatical en el sujeto con un porcentaje de 2% en ambas L1. 3) La falsa selección de la concordancia de género gramatical en la frase nominal del atributo con un 3% en L1 inglés y un 1% en L1 francés y, finalmente, 4) la falsa selección de la concordancia gramatical entre el sujeto y el verbo *ser* que en ambas L1 corresponde al 6%.

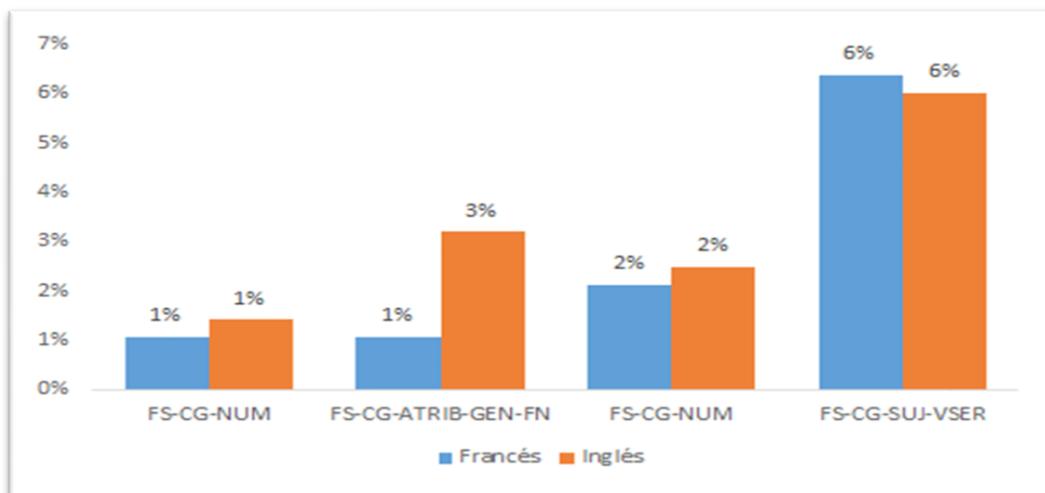


Gráfico 4. Similitudes en las frecuencias de errores entre los francófonos y anglosajones

Esta distribución de los errores con menor frecuencia revela una menor dificultad para este tipo de concordancias y un uso con mayor precisión. Estos hallazgos suponen un avance importante, dado que permiten identificar cuáles son los errores en los que debería centrarse el análisis respecto de la conformación de la interlengua. Es decir, la concordancia gramatical de género y número en el atributo prueban ser una dificultad mayor dadas las características morfológicas de los sustantivos y adjetivos en español, así también la distancia lineal entre los elementos entre los cuales debe darse la concordancia gramatical.

5. CONCLUSIONES

En este artículo, hemos presentado un estudio descriptivo con un enfoque de análisis de datos mixto con el propósito de explorar en dos subcorpus del corpus de aprendientes de Español como Lengua Extranjera, CAELE, las dificultades que se observan en aprendientes anglosajones y francófonos de nivel A2 con respecto al uso del verbo *ser* en las estructuras de atributos nominales en un contexto de tareas de escritura académica.

El CAELE se constituye en un corpus robusto de 1217 textos que permite examinar distintas problemáticas lingüísticas y realizar análisis de corpus a nivel léxico-gramatical. Este importante recurso lingüístico y su metodología de trabajo facilita realizar análisis de muestras textuales auténticas y representativas de aprendientes de ELE. La investigación realizada a partir del CAELE aporta información relevante sobre los procesos que sigue el aprendiente hasta alcanzar un determinado dominio en las estructuras de concordancia gramatical. Como se ha señalado, los estudios de corpus de aprendientes en formato computacional han contribuido con información relevante sobre las fases por las que pasa el aprendiente hasta la completa adquisición y las características de su interlengua en función del nivel de dominio en el que se encuentre. CAELE se constituye en un inventario abierto que año tras año considera la recolección y el procesamiento de nuevos textos.

Los resultados de la investigación han arrojado que para los aprendientes anglosajones hay mayores dificultades en la selección de las categorías de género y número en los adjetivos, especialmente en la concordancia de los atributos. En cambio, para los aprendientes del francés, si bien presentan un número de errores en dichos aspectos gramaticales, su frecuencia es notoriamente menor.

La interpretación de las posibles explicaciones de por qué los aprendientes de ambas L1 cometieron los errores está supeditada tanto al sistema lingüístico de su L1 como a la conformación y evolución de la interlengua en ELE. Para el caso del francés, puede concluirse que las estructuras erróneas podrían atribuirse a tres estrategias: (1) la transferencia de la L1 respecto de la semejanza en ambas lenguas dado su origen romance, (2) a la confusión que podría generar las reglas entre el español y el francés similares, como en el caso del género morfológico de los sustantivos y su correlación con los adjetivos en español. En este último se infiere que, a pesar de las similitudes entre el género del español y el francés, muchos de estos son diferentes en cada lengua, y (3) la posibilidad de la distancia entre los elementos que concuerdan tanto en el sintagma nominal como en el atributo, es decir, "los elementos de concordancia pueden situarse muy cercanos entre sí, como un sustantivo y su determinante definido o indefinido, o muy lejanos, tal como el sustantivo y su adjetivo" (Paquet, 2017: 60). Este distanciamiento podría afectar en la concordancia, considerando que muchas de las oraciones se apoyan en elementos modificadores de los sustantivos provocando un distanciamiento que conduciría al error.

Para el caso del inglés, los errores podrían atribuirse a tres estrategias: (1) la transferencia directa negativa dada la distancia entre las lenguas (ausencia de género y número gramatical de los adjetivos y sustantivos en casos específicos), (2) la sobregeneralización de reglas de fácil acceso para el aprendiente y (3) la reducción de estructuras complejas a simples como en el caso de la concordancia entre el sujeto y el verbo copulativo *ser*. En estas tres estrategias, la que mayor incidencia tiene es la transferencia negativa, dado que entre el inglés y el español muchas reglas reflejan la ausencia de rasgos morfológicos y gramaticales, la ausencia en el uso de elementos que faciliten la concordancia gramatical (como artículos o modificadores del sustantivo, donde el género garantiza la cohesión sintáctica en la oración (Paquet, 2017)) y las diferencias entre la organización interna de los sintagmas nominales y verbales. En cuanto a la sobregeneralización, es común observar que los aprendientes de inglés suelen generalizar reglas del género que resultan simples a la vista, dadas las terminaciones en la vocal "o" y la vocal "a" evocando la primera al género masculino y la segunda al femenino. Sin embargo, en las palabras con género neutro, se observa una tendencia a masculinizar las palabras terminadas en "e". Respecto de estudios que corroboran la dificultad que presenta aprender una LE a partir de una L1 sin género (Keating, 2009), puede adquirirse la asociación de elementos y la concordancia gramatical, sin embargo, no será procesada de la misma forma que un hablante nativo. También afirman que la influencia de la distancia entre los elementos podría afectar a la comprensión y sensibilidad al momento de aprender las reglas de la concordancia y, por ende, sería un proceso más lento y tardío que la adquisición de otras reglas gramaticales (Keating, 2009; Paquet, 2017).

Por lo tanto, la consideración de estas diferencias o similitudes en ambas lenguas es preponderante al momento de diseñar el proceso de aprendizaje en estas estructuras del verbo copulativo *ser*. Esto debido a que, en consonancia con los errores identificados, se pueda considerar la distancia y similitudes en ambas lenguas. Asimismo, puede arrojar luces sobre cómo la organización de las estructuras de las oraciones copulativas que son base para la precisión gramatical (concordancia entre el sujeto y el verbo, concordancia de género y número) frente al aprendizaje de la concordancia gramatical, es decir, cuáles de estos podrían priorizarse para la enseñanza y así cautelar un adecuado desarrollo de las estructuras más complejas del español respecto de las lenguas en estudio para esta investigación.

Finalmente, el proceso de enseñanza-aprendizaje en torno a la concordancia gramatical presente en las oraciones con el verbo *ser* es una temática morfosintáctica muy relevante en ELE y genera un impacto en la producción de las tareas de escritura académica por parte de los aprendientes francófonos y anglosajones. En el contexto del aula, la concordancia debe articularse con actividades que incorporen las cuatro habilidades comunicativas y ejemplos auténticos que reflejen la lengua en uso. Dichos ejemplos permiten al docente proporcionar explicación gramatical y a los aprendientes de ELE de competencia lingüística de nivel A2 visualizar, reflexionar y corregir los usos incorrectos.

AFILIACIÓN DEL PROYECTO

Este artículo se ha desarrollado en el contexto del Proyecto Fondecyt 1180974 "Diseño e implementación de un corpus escrito de aprendientes de Español como Lengua Extranjera (ELE) para el análisis de la interlengua", 2018-2021.

NOTAS

1 En el corpus CAELE cada texto está identificado con una etiqueta del tipo S43A2T4I2016, en la cual S43 corresponde al número del sujeto de la muestra, A2 al nivel de competencia; T4 al número del texto escrito por el sujeto; I corresponde a la L1 inglés y finalmente 2016 al año en que se recolectó el texto.

REFERENCIAS BIBLIOGRÁFICAS

- Agudelo, S. (2018). La interlengua de los estudiantes francófonos de ELE: el caso de ser y estar. *Forma y función*, 31(2), pp. 93-128.
- Aránzazu, M. (2004). La atención a la forma aplicada a la adquisición de la concordancia del género de los sustantivos "problemáticos" en español. *Interlingüística*, (15), pp. 1217-1224.
- Baralo, M. (1996). Adquisición y/o aprendizaje del español/LE. En: Tendencias actuales en la enseñanza del español como lengua extranjera II. Actas del VI Congreso Internacional de ASELE. [en línea] León: Universidad de León, págs. 63-67. Disponible en: https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/06/06_0062.pdf [Consultado el 30 de julio de 2020].
- Baralo, M. (1998). Ser y estar en los procesos de adquisición de lengua materna (lm) y de lengua extranjera (le). En: Español como Lengua Extranjera: Enfoque Comunicativo y Gramática. Actas del IX Congreso Internacional de ASELE. [en línea] Santiago de Compostela: Universidade de Santiago de Compostela, págs. 291-300. Disponible en: https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/09/09_0294.pdf [Consultado el 10 de agosto de 2020].
- Bautista, M. (2019). *Terminología y ontologías: un estudio alemán-inglés-español basado en corpus*. Granada: Editorial Comares.
- Bond, K, Gabriele A, Fiorentino, R. y Alemán Bañón, J. (2011). Individual differences and the role of the L1 in L2 processing: An ERP investigation. In: Herschensohn J and Tanner D (eds) Proceedings of the 11th Generative Approaches to Second Language Acquisition. Somerville, MA: Cascadilla Proceedings Project, 17-29.
- Buyse, K. (2011). ¿Qué corpus utilizar para qué fines en la clase de ELE? En: Del texto a la lengua: la aplicación de los textos a la enseñanza-aprendizaje del español L2-LE. Ponencia llevada a cabo en el XXI Congreso Internacional de ASELE. [en línea] Salamanca: págs. 277-288. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=5419362> [Consultado el 27 de agosto de 2020].
- Campillos Llanos, L. (2014). Errores léxicos en el español oral no nativo: análisis de la interlengua basado en corpus. *Revista ELUA*, Alicante, v. 28, pp.85-124.
- Cruz Piñol, M. (2012). *Lingüística de corpus y enseñanza del español como 2/L*. Madrid: Arco Libros.
- Doquin de Saint Preux, A. y Sáez Garcerán, P. (2014). Errores gramaticales persistentes en la interlengua de estudiantes francófonos de ELE de nivel superior. *MARCO ELE Revista de Didáctica*, 19, 1, pp.-19.
- Elejalde, J. y Ferreira, A. (2016). Errores de transferencia en comunidades de aprendizaje en línea por aprendientes de español como lengua extranjera (ele). *Trabalhos em Linguística Aplicada*, 55(3), pp. 619-650.
- Fernández, S. (1997). *Interlengua y análisis de errores*. Madrid: Edelsa.
- Ferreira, A. (2016). *Prueba de multinivel con fines específicos académicos*. Universidad de Concepción, Chile.

- Ferreira, A. y Elejalde, J. (2017). Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE. *Revista Brasileira de Lingüística Aplicada*, 17(3), pp. 509-538. <https://doi.org/10.1590/1984-6398201710927>
- Ferreira, A., y Elejalde, J. (2019). Hacia un perfil lingüístico-comunicativo del estudiante de Español como Lengua Extranjera para fines Académicos. *Revista Nebrija De Lingüística Aplicada a La Enseñanza De Lenguas*, 13(27), 145-165. <https://doi.org/10.26378/rnlael1327326>
- Ferreira, A., y Elejalde, J. (2020). Propuesta de una taxonomía etiológica para etiquetar errores de interlengua en el contexto de un corpus escrito de aprendientes de ele. *Forma y Función*, 33(1), 115-146 <https://revistas.unal.edu.co/index.php/formayfuncion/article/view/84182>
- Ferreira, A. y Elejalde, J. (2020). Análisis de Errores en ELE por Transferencia Negativa del inglés como L1 en clases presenciales versus clases no presenciales en comunidades virtuales. *Revista Brasileira de Linguística Aplicada*, 20(3), pp. 591-626. <https://doi.org/10.1590/1984-6398201710927>
- Ferreira, A. (2018-2021). FONDECYT- CONICYT N° 1180974. "Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua". Proyecto de Investigación. Conicyt, Chile.
- Fleissig, D. (2009). *El análisis de los errores de la concordancia de los estudiantes universitarios anglohablantes*. Tesis de Licenciatura. Universidad de Michigan.
- Gilquin, G. (2008). Combining contrastive and interlanguage analysis to apprehend transfer: detection, explanation, evaluation. In: G. Gilquin, S. Papp y M^a. Díez-Bedmar, eds., *Linking up contrastive and learner corpus research*, 1^a ed. Amsterdam/New York, NY: Rodopi. Language and Computers - Studies in Practical Linguistics, págs. 3-33.
- Granger, S. (1996). From CA to CIA and back. An integrated approach to computerized bilingual and learner corpora. In: K. Aijmer, B. Altenberg, y M. Johansson, eds., *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies*, 1^a ed. Cambridge: University Press, págs. 37-51
- Granger, S. (2002). A bird's-eye view of learner corpus research. In: S. Granger, J. Hungand y S. Petch-Tyson, eds., *Computer learner corpora, second language acquisition and foreign language teaching*, 1^a ed, Philadelphia: EE.UU: John Benjamins, págs. 3-33.
- Granger, S., Hung, J. y Petch-Tyson, S. (2002). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam: John Benjamins.
- Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, 20(3), pp. 465-480.
- Granger, S. (2004). Computer learner corpus research: Current status and future prospects. *Language and Computers*, 52(1), pp. 123-145.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching. In: Aijmer, K., eds., *Corpora and language teaching*, 1^a ed. Amsterdam: John Benjamins, págs. 13-32.
- Granger, S. (2012). How to use foreign and second language learner corpora. In: A. Mackey y S. Gass. eds., *Research methods in Second Language Acquisition: A practical guide*, 1^a ed. Malden: Blackwell Publishing, págs. 7-29.
- Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), pp. 7-24.
- Guijarro-Fuentes, P. (2008). ¿Ser o estar? Cuestiones sin resolver para el español como lengua extranjera. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 2(3), pp. 44-55.
- Hasselgard, H. y Johansson, S. (2011). Learner corpora and contrastive interlanguage analysis. In: F. Meunier, S. De Cock, G. Gilquin y M. Paquot, eds., *A taste for corpora: In honour to Sylviane Granger*, 1^a ed. Amsterdam/Philadelphia: John Benjamins Publishing Company, págs. 33-62.

- Ishikawa, K. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In: S. Ishikawa, Ed., *Learner corpus studies in Asia and the world*, 1ª ed. Kobe: Kobe University, págs. 91-118.
- Keating, G.D. (2009). Sensitive to violations of gender agreement in native and no native Spanish: An eye-movement investigation. *Language Learning*, 53(3), pp. 503-535.
- Kilgarriff, D., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. y Suchomel, V. (2014). *The Sketch Engine: Ten years on. Lexicography*, 1: 7-36.
- Kilgarriff, D., Rychlý, P., Smrž, P. y Tugwell, D. (2004). *The Sketch Engine. Information Technology*. 105, 116.
- Liceras, J. (1983). *Contrastive analysis and the acquisition of Spanish syntax by English speakers*. Tesis Doctoral. Universidad de Toronto.
- Lozano, C. (2015). Learner corpora as a research tool for the investigation of lexical competence in L2 Spanish. *Journal of Spanish Language Teaching*, 2(2), pp. 180-193.
- Méndez de Sifontes, I. y Rojas, I. (2013). An error analysis of persistent grammatical errors carried over from beginner to intermediate Spanish. *Revista Electrónica de Lingüística Aplicada*, 12, pp. 63-85.
- Montes, J. (1997). Notas sobre el género gramatical en Bello y en la actualidad. *Thesaurus. Boletín del instituto Caro y Cuervo*, 52(1-3), pp.144-155.
- Naranjo, J. (2009). *Análisis de errores en la interlengua de estudiantes de español como lengua extranjera. Un estudio sobre la correlación de los errores con la comunicabilidad en la producción escrita de estudiantes de habla inglesa*. Tesis de magister. Universidad de Bergen.
- Paquet, P. (2017). *Influencia de las propiedades de la L1 y del nivel de competencia en la adquisición de la concordancia de género: una comparación entre estudiantes francófonos y anglófonos de español L2*. Tesis de Doctorado. Universidad de Alicante.
- Pastor Cesteros, S. (2001). La concordancia en la interlengua de los aprendices de español como lengua extranjera. *ELUA. Estudios de Lingüística, Anexo 1*, pp. 95-125.
- Pastor Cesteros, S. (2004). *Aprendizaje de segundas lenguas: Lingüística aplicada a la enseñanza de idiomas*. Alicante: Editorial Universidad de Alicante.
- Penadés, I. (1999). *Lingüística contrastiva y análisis de errores*. Madrid: Edinumen.
- Prieto, C. (2011). *Análisis de errores de estudiantes anglófonos*. Memoria de investigación. Universidad de Alcalá.
- Rakaseder, J. y Schmidhofer, A. (2014). Errores de producción de textos por parte de germano parlantes en las primeras etapas de aprendizaje de ele. *MARCO ELE Revista de Didáctica*, 19, pp. 1-17.
- Real Academia Española. (2005). *Diccionario Panhispánico de dudas*, 1ª ed. Madrid: Taurus.
- Real Academia Española. (2010). *Manual de la Nueva Gramática de la Lengua Española*. Buenos Aires: Espasa.
- Real Academia Española. (2014). *Diccionario de la lengua española*, 23ª ed. Buenos Aires: Espasa.
- Rubinjon, L. (2014). Análisis de los errores de concordancia en la expresión escrita de los aprendientes serbios de ELE. *Linred: Lingüística en la Red*, 11, pp. 1-34.
- Sánchez, A. (2015). La Investigación de Corpus de Aprendientes y el desarrollo de los estudios de la interlengua del español. *Language design: Journal of theoretical and experimental linguistics*, 17, pp. 57-83.
- Santos Gargallo, I. (1993). *Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva*. Madrid: Síntesis.
- Vázquez, G. (1991). *Análisis de errores y aprendizaje de español/lengua extranjera*. Frankfurt am Main: Peter Lang.