Errores sistemáticos en un corpus de aprendices de ELE con implicaciones para el feedback correctivo escrito enfocado[1]

Anita Ferreira Cabrera
Universidad de Concepción

1. Introducción

El corpus de aprendices de ELE, CAELE, es una colección de 418 textos escritos por 62 aprendices de ELE. Estos textos han sido recolectados durante tres procesos de producción escrita entre los años 2014 y 2015 y están almacenados y procesados en formato digital. El objetivo principal de este estudio es determinar los errores lingüísticos más frecuentes y recurrentes en dicho Corpus con la finalidad de apoyar la selección de errores adecuados para los procesos de corrección gramatical a través de estrategias de Feedback Correctivo Escrito (FCE) enfocado.

En el FCE enfocado se selecciona un número de errores específicos para ser corregidos y se descartan otros. Según Ellis, Sheen, Murakami y Takashima (2008), existen razones teóricas sólidas que avalan el argumento de que el FCE enfocado puede ser más acertado que el FCE no enfocado, en el cual se corrigen todos los errores del escrito. Las causas o razones se deben a que los estudiantes son más sensibles en poner atención a correcciones dirigidas a uno o a un número limitado de errores para desarrollar mayor comprensión de la naturaleza del error.

Los resultados de esta investigación evidencian que los dos tipos de errores más *frecuentes* y recurrentes en el corpus CAELE corresponden a la *falsa selección del género gramatical* y a *la omisión de la tilde en las palabras.* Estas tendencias tendrán implicaciones para el mejoramiento de la destreza escrita a través de la focalización de errores recurrentes para su tratamiento por medio de estrategias de FCE enfocado.

El corpus de aprendices se constituye en un valioso recurso para la investigación en el ámbito de la ASL y en la enseñanza de L2. Un corpus electrónico de aprendices de lengua (del inglés *Computer Learner Corpora, (CLC)*) es la colección de datos lingüísticos auténticos (textos orales o escritos) donde se constata el uso de la lengua objeto de estudio (segunda lengua, (L2) o lengua extranjera (LE)) (Granger, 2003, 2004). Estas colecciones de textos pueden ser posteriormente procesadas por medio de herramientas de software especializado para la recuperación de información acorde con determinados criterios de búsqueda. En este contexto, si bien uno de los objetivos principales en la ASL es construir modelos de las representaciones subyacentes al proceso de aprendizaje de un estadio en particular durante la adquisición de una L2, los métodos de recolección se han realizado a través de estudios experimentales, observacionales o métodos introspectivos. Así como lo señala Granger (2004) la investigación en ASL ha

1. Este estudio se ha desarrollado en el contexto del proyecto de investigación CONICYT-FONDECYT No. 1140651 "El Feedback Correctivo Escrito Directo e Indirecto en la adquisición y aprendizaje del Español como Lengua Extranjera".

favorecido los métodos con variables controladas y experimentales, relegando los datos lingüísticos del uso natural de la lengua.

2. Sustentos teóricos

El Análisis de Errores Asistido por Computador (del inglés *Computer-aided-error-analysis (CEA))* es un enfoque de investigación basado en corpus electrónico de aprendices (del inglés *Computer Learner Corpora (CLC*)) y en LC para la identificación, clasificación y explicación de los errores de forma digital (Dagneaux, Denness y Granger, 1998; Piñol, 2012). Este tipo de estudio permite la posibilidad de procesar los errores de lengua y su frecuencia en el contexto de un texto (oral o escrito) a través de herramientas de análisis computacional (Granger, 2004). La metodología de análisis de errores asistido por computador se basa en los procedimientos establecidos por Corder (1975): la identificación, la descripción, la clasificación, la explicación y la evaluación de la gravedad del error (Corder, 1975; Ellis, 1997; Ferreira, 2003; Ferreira, Elejalde y Vine, 2014; Torijano, 2006). Junto con dichos procedimientos, se realiza un etiquetado semiautomático con software especializado para etiquetaje de corpus, acorde con un sistema de anotación previo. Finalizado el etiquetaje, se realiza el procesamiento de los datos para explorar el corpus y obtener los resultados según las variables del estudio.

2.1. Corpus de Aprendices de ELE

En el ámbito de la lingüística aplicada a la enseñanza y aprendizaje del español como LE o L2, existen corpus orales y escritos recolectados en formato electrónico. La mayoría de estos corpus consideran variables como el nivel de competencia, la lengua materna y años de estudio de los aprendientes con el objeto de identificar aspectos diferenciadores del proceso de adquisición. Algunos de estos corpus son, por ejemplo, el Corpus Escrito del Español como L2 (CEDEL2) de aprendices de inglés como L1. Este corpus considera producciones textuales en todos los niveles de competencia y se recolectó con fines contrastivos entre el español y el inglés (Lozano y Mendikoetxea, 2013). CEDEL2 está conformado por alrededor de 750.000 palabras en formato electrónico y se ha recolectado de forma sistemática vía online con estudiantes y profesores voluntarios. El corpus está disponible para ser utilizado como fuente de datos para los investigadores de ELE y con abundancia de ejemplos para la enseñanza de ELE.

Otro corpus interesante es el Corpus de Aprendices de Español como Lengua Extranjera (CAES) (Rojo y Palacios, 2015). Corresponde a un conjunto de textos escritos producidos por estudiantes de español como LE de todos los niveles de competencia. Este corpus ha sido elaborado en colaboración por investigadores y docentes de la Universidad de Santiago de Compostela y del Instituto de Cervantes, España. Los textos han sido generados por aprendientes de diferentes lenguas maternas: árabe, chino, mandarín, francés, inglés, portugués y ruso. El total de textos recolectados es de 3.878 provenientes de 1.423 estudiantes que escribieron dos o tres textos cada uno.

2.2. La frecuencia y la sistematicidad de los errores

Los errores se identifican acorde con el número de veces que aparecen, el contexto y la repetición de estos. Para ello, existen dos formas de precisar la identificación de los errores en valor numérico:

(1) *la frecuencia*, entendida como el número de veces que ocurre un error en cualquier muestra de textos orales o escritos. Esta frecuencia, en términos generales, describe la interlengua de la muestra que se observa y puede determinarse según dos modalidades, la absoluta y la relativa. La primera es el conteo por cada error y su aparición, a diferencia de la segunda que corresponde al número de ocurrencias en el que aparece un error en proporción con la extensión de un texto. El resultado será un criterio numérico útil para determinar la gravedad de los errores en relación con la producción escrita u oral.

(2) *La sistematicidad*, a la hora de valorar los errores, la sistematicidad está relacionada con qué tipo de errores son sistemáticos y repetitivos dentro de la misma producción textual de un sujeto. Este término permite evaluar la gravedad del error para determinar si son desviaciones sistemáticas, lapsus o equivocaciones. Este tipo de medición permite registrar los errores reiterativos en distintas producciones de un mismo sujeto. El análisis de recurrencia de los errores de un corpus contribuye a identificar qué errores aún persisten incluso durante o después de la instrucción formal en la clase de ELE. En este sentido, el procedimiento metodológico de este tipo de análisis permite discernir si un estudiante o grupo domina o no una regla.

3. El estudio: errores frecuentes y sistemáticos en el CAELE

El diseño de investigación de este estudio es descriptivo y longitudinal con un enfoque de análisis de datos cuantitativo. El objetivo principal es determinar los errores más frecuentes y sistemáticos en el Corpus CAELE.

3.1. Muestra de Aprendientes de ELE

La muestra se constituyó por un total de 62 estudiantes universitarios extranjeros de una universidad chilena (26 de nivel A2+ y 36 de nivel B1), correspondientes a la cohorte 2014 (15 sujetos= 24%) y 2015 (47 sujetos= 76%) de cursos de español general como lengua extranjera. Las lenguas maternas de los sujetos corresponden al alemán con 20 sujetos (32%), francés 17 (27%), inglés 17 (27%), portugués 2 (3%), sueco 2 (3%), checo 2 (3%), italiano 1 (2%) y ruso 1 (2%). Como se observa en la Tabla No.1, cada sujeto escribió un total de 6 textos en la cohorte 2014 y 7 textos en la cohorte 2015. El número de textos producidos en el nivel A2 corresponde a un total de 176 textos (42%) y a 242 textos (58 %) en el nivel B1. Cada texto se generó en el marco del desarrollo normal de las clases de ELE.

Sujetos por nivel		No. Sujetos	Total sujetos	No. Textos	Total textos	%=418
A2+	2014	5	26	29	176	42%
	2015	21		147		

B1	2014	10	36	60	242	58%
	2015	26		182		
Totales		**62**	**62**	**418**	**418**	**100%**

TABLA No. 1 Muestra Corpus CAELE

3.2. El Corpus de Aprendices de ELE

El CAELE es una colección de 408 textos digitales de aprendices de español como lengua extranjera generados en el contexto de los proyectos de investigación FONDECYT No. 1110812 y No. 1140651 (Ferreira, 2011 y 2014). Estos textos fueron recolectados durante tres periodos de clases de ELE entre los años 2014 y 2015. Los sujetos de la cohorte 2014 escribieron 6 textos cada uno con un total de 89 textos. La cohorte 2015 escribió 7 textos por cada sujeto con un total de 329 textos. De los 418 textos, los aprendientes del nivel A2+ produjeron un total de 176 (42%) textos y los del nivel B1 242 (58%). La modalidad textual predominante es la narrativa.

3.3. Metodología de la investigación

Para la identificación, clasificación y anotación de los errores, se procedió con los criterios y taxonomía de errores del modelo metodológico desarrollado por Ferreira (2011 y 2014). El procesamiento de los datos se realizó a través del software *Uam Corpus Tool* versión 3.2[2].

3.3.1. Recopilación del corpus

En la recolección del corpus CAELE se cauteló que los textos fueran auténticos y pertinentes para la investigación. Dichos textos fueron escritos en un período de tres meses con un total de 12 semanas en el desarrollo normal de las clases presenciales de los cursos de ELE. Los aprendices escribieron en promedio un texto por semana. La extensión de los textos fue de 150 a 200 palabras en el nivel A2 y de 250 a 300 en el B1. Una vez finalizada la escritura, cada texto fue guardado en formato *.txt* con codificación UTF-8*,* ordenada y compilada para procesarlos en el software *Uam Corpus Tool* versión 3.2.

3.3.2. Identificación y clasificación de los errores

Como se observa en la Tabla No.2, los niveles de categorización en la taxonomía de los errores corresponden a: gramática (categorías gramaticales y coherencia textual), léxico (a nivel morfológico y de significado) y ortografía (acentual, literal, dierética y diacrítica). El criterio lingüístico se fundamenta en la visión del error como un proceso sistemático que todo aprendiente de lenguas experimenta durante su aprendizaje de la L2 (Ellis, 1997). Autoras como Vázquez (2009), Quiñones (2009a, 2009b) y Alexopoulou (2006)

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

han desarrollado taxonomías para ELE basadas en el sistema lingüístico del español, para dar cuenta de los errores por cada categoría y nivel de escritura (palabra u oración).

La anotación de los errores se basó en la identificación de las modificaciones y/o usos de la L2, las cuales clasifican el error en: (1) omisión de elementos oracionales o categorías gramaticales (OMI: "..._____ * *Gente siempre hablan rapido*..."), (2) falsa selección (FS: "...*voy* ***con** bus a Viña*..") de cualquier elemento dentro del sintagma nominal o verbal, (3) forma errónea (FE: "... *esto pasa* ***por qué** no conoce la biblioteca*...") en la elección de la forma acorde con el significado, (4) adición de elementos gramaticales (ADI: "...*he decidido* ***de** aprender*..."), y (5) elección errónea de la realización semántica de una palabra o la ortografía de una palabra en específico (EE: "...*los* ***musicos** chilenos son importante para la gente*...").

Criterio	Nivel	Categoría	Subcategoría	Tipo		Descripción del error
Lingüístico	Palabra y oración	Gramática	Categorías gramaticales	Preposición		ADI (adición) FS (falsa selección) OMI (omisión) FE (forma errónea) EE (elección errónea)
				Artículo		
				Pronombre		
				Verbo		
				Adjetivo		
				Sustantivo		
				Adverbio		
			Coherencia textual	Concor. Sintáctica	Nominal	
					Verbal	
					Atributo	
				Puntuación		
				Conectores		
	Palabra	Léxico	Léxico creado por derivación			
			Léxico morfología	Verbal		
				Sustantivo-adjetivo		
				Derivación		
			Léxico significado	Forma errónea		
				Elección errónea significado		
				Léxico inncesario		
		Ortografía	Acentual			
			Literal			
			Dierética			
			Diacrítica			

TABLA No. 2: Taxonomía para la identificación y clasificación de errores, Proyecto FONDECYT No.1140651 (Ferreira, 2014)

3.3.3. Etiquetado de errores en el software Uam Corpus Tool

Para el etiquetamiento de errores se utilizó el programa especializado en anotación de corpus desarrollado por Mick O´Donell (2008) en el contexto del Ministerio de Educación y Ciencia de España bajo el número de licencia HUM2005-01728/FILO (The WOSLAC Project).

El software cuenta con herramientas para el etiquetado manual, semiautomático o automático. Además permite anotar los errores a partir de un esquema de anotación de forma manual y asignarle atributos al texto para obtener resultados acordes con las variables del estudio (lengua materna, nivel, entre otros).

3.3.4. Procesamiento de los datos: obtención de las matrices para el análisis

Para determinar los errores más recurrentes del corpus CAELE, se obtuvo la frecuencia de los errores en toda la muestra por cada sujeto acorde con el nivel de competencia. Para ello, el programa *Uam Corpus Tool*, ofrece la opción de procesar los datos a través de matrices en tres formatos: 1) descripción general del corpus *(describe a dataset)*, 2) descripción del corpus comparando dos o más variables *(compare several datasets)* y 3) descripción de cada archivo digital del corpus *(describe each file)*.

Para contar los errores recurrentes se procedió de acuerdo con el criterio de determinar cuántos errores de "x categoría gramatical o tipo de error" aparecían en "x textos". De esta forma, se cauteló la identificación de todas las recurrencias posibles en los textos por sujeto y la clasificación de los errores más recurrentes en toda la muestra. La Tabla No. 3 muestra un ejemplo de la identificación de los errores cometidos en cada texto, su frecuencia y marcaje por recurrencia. Para ello, el tipo de error se marcó con sombreado gris las veces en las que el sujeto incurrió en un error a través de los textos. Por ejemplo, en el error "*omi-prep-de*" el sujeto incurrió en los textos 2, 3, 4, 5, 6 (un total de 5 textos), con una frecuencia entre una o dos veces de aparición por cada texto. Esto indica una tendencia del error que se mantiene en el tiempo de manera consistente. También puede verse que en el error "*omi-art-def*", el sujeto sólo incurrió en tres textos, sugiriendo que el error no es consistente y aparece menos constante.

Terminado este primer proceso los datos fueron cuantificados en una tabla Excel para obtener la frecuencia absoluta y los porcentajes relativos con respecto a la recurrencia por nivel de competencia. Sobre la base de estos datos, se determinó que la recurrencia sería clasificada a partir de un número mínimo de cuatro sujetos por nivel.

Análisis de recuerrencia nivel A2+ por sujeto						
Sujeto: 8 L1: alemán						
Tipo de error	conteo errores por recurrencia					
ARTÍCULO DEFINIDO	Texto 1	Texto 2	Texto 3	Texto 4	Texto 5	Texto 6
omi-art-def	0	1	0	0	2	1
falsa-s-art-def	0	0	0	0		0
adi-art-def	0	0	0	0		0
PREPOSICIONES						
omi-prep-a	0	0	0	1	1	0
omi-prep-de	0	2	1	1	1	1
VERBO-MODO						
falsa-s-modo-subj	0	0	3	1	2	0
falsa-s-modo-indic	0	1	0	1	1	0
ORTOGRAFÍA						
omi-orto-esdrújula	4	2	1	3	1	1
ee-orto-acen-esdrújula	0	0	0	1	1	0

TABLA No. 3: Matriz por sujeto y conteo de errores por recurrencia.

4. Resultados

4.1. Frecuencias de errores en el Corpus CAELE

El total de errores identificados según el criterio lingüístico fue de 8.731 (ver en Tabla 5), los cuales se distribuyen en cuatro categorías con sus respectivas subcategorías: 1) gramática con un 29% (2.552 errores clasificados en omisión, falsa selección y adición de categorías gramaticales), 2) la coherencia textual con un 18% (1.569 errores clasificados en la falsa selección de la concordancia gramatical), 3) la ortografía con un 39% (3.448 errores, distribuidos en la omisión y colocación errónea de la acentuación) y 4) el léxico con un 13% (1.162 errores, clasificados en aspectos morfológicos y semánticos).

En el contexto de los errores gramaticales, en lo referido específicamente a las *categorías gramaticales,* se puede señalar que la mayor problemática la presentan las preposiciones con 989 errores (ver Tabla No. 4), lo que representa un 39% del total de 2.552 errores. Esto muestra que en general la complejidad del sistema preposicional representa un área de dificultad mayor para los aprendices de ELE corroborándose con esto los resultados de investigaciones previas (Campillos, 2014; Ferreira y Lafleur, 2015).

Con respecto a los errores en la subcategoría de la *coherencia textual*, la falsa selección de la concordancia gramatical de género es el tipo de mayor frecuencia con 968 errores que corresponden a un 62%; esto se puede explicar por el hecho de que las reglas de concordancia en muchos casos los aprendices de ELE las generalizan o las desconocen.

En la *ortografí*a acentual, las palabras esdrújulas presentan un total de 823 errores, cuya problemática responde a la omisión y la elección errónea de la tilde. En este as-

pecto, la ortografía puede verse influenciada por varios factores como, por ejemplo, la influencia de la lengua materna, el desconocimiento de la regla o la sobregeneralización de las normas de tildación.

Tipo de errores				2014	2015-1	2015-2	SUBTOTAL	%=subtotal	%=8731
Criterio Lingüístico	Gramática	Categorías gramaticales	Preposición	246	293	450	989	39%	29%
			Verbo	112	219	300	631	25%	
			Artículo	73	216	206	495	19%	
			Pronombres	46	106	133	285	11%	
			Adjetivo	19	18	47	84	3%	
			Adverbio	4	20	13	37	1%	
			Sustantivo	3	17	11	31	1%	
	SUBTOTAL			503	889	1160	2552	100%	
	Coherencia textual	Concordancia sintáctica	Género	244	186	538	968	62%	18%
			Número	78	67	197	342	22%	
			Suj-pred	49	101	109	259	17%	
	SUBTOTAL			371	354	844	1569	100%	
	Ortografía	Acentual	Esdrújula	134	248	441	823	24%	39%
			Aguda	146	214	334	694	20%	
			Grave	41	85	131	257	7%	
			Sobreesdrújula	2	12	13	27	1%	
		Dierética		91	150	251	492	14%	
		Diacrítica		100	153	301	554	16%	
		Literal		86	148	367	601	17%	
	SUBTOTAL			600	1010	1838	3448	100%	
	Léxico	Morfología	Sust-adj-adv	91	166	194	451	39%	13%
			Verbo	37	73	125	235	20%	
			Fs-morf-gén	13	33	66	112	10%	
			Fs-morf-num	0	0	8	8	1%	
		Forma errónea		14	13	48	75	6%	
		Falsa selección léxica		33	103	145	281	24%	
	SUBTOTAL			188	388	586	1162	100%	
TOTAL				1662	2641	4428	8731	100%	

TABLA No.4: Resultados generales del corpus CAELE: criterio lingüístico

4.2. Recurrencia de errores en el Corpus CAELE

En relación con la sistematicidad de los errores, se procedió a realizar un análisis de la recurrencia de los errores; para ello se procedió al conteo de los errores que se observan en varios textos producidos en distintos momentos por un mismo sujeto. Además, en esta investigación para la delimitación de la recurrencia de errores se consideró un número mínimo de cuatro sujetos para cada problemática, puesto que así se cautela la representatividad de dicha problemática. Como se puede ver en la Tabla No. 5 el número total de errores recurrentes producidos por un mínimo de 4 sujetos corresponde a 2.412 errores. Los resultados muestran que la ortografía presenta una recurrencia mayor con un 57% del total de errores, seguida por la concordancia sintáctica con un 24%, la preposición con un 10%, el artículo con un 6% y el verbo con un 2,5%.

Estos resultados se asemejan a algunos resultados de la frecuencia general, específicamente en lo referido a los errores de ortografía acentual y de la concordancia gramatical de género. En consecuencia, estas tendencias representan las áreas más problemáticas que no sólo son frecuentes, sino que además se mantienen en el tiempo. Asimismo, la recurrencia de estos errores indica que corresponde a un tipo de error sistemático cuya interferencia en el proceso de aprendizaje dependerá, principalmen-

te, de la gravedad del error en relación con los aspectos comunicativos (incomprensión del mensaje) y lingüísticos de la L2.

Recurrencia por categoría general						
Categoría	**Tipo de error**	**Fr. Abs**	**%=2412**		**Textos**	**No. Sujetos**
Ortografía	omi-orto-esdrújula	490	20%	57%	6;5;4;3	38 de 62
	omi-orto-aguda	296	12%		6;5;4	22 de 62
	omi-orto-grave	20	1%		2	8 de 62
	ee-orto-acento-grave	64	3%		3;2	15 de 62
	omi-orto-hiato	226	9%		5;4;3;2	35 de 62
	omi-orto-diacrítica	219	9%		6;4;3;2	30 de 62
	ee-orto-diacrítica	7	0%		2	6 de 62
	ee-orto-lit-s	9	0%		2	4 de 62
	ee-orto-lit-z	9	0%		2	4 de 62
	omi-orto-mayúscula-inicial	31	1%		2	8 de 62
	fs-orto-mayús-inicial	12	0%		2	4 de 62
Concordancia sintáctica	fs-cg-gen	270	11%	24%	6;5;4;3	27 de 62
	fs-cg-gen-atributo	40	2%		2;4	12 de 62
	fs-cg-gen-epiceno	11	0%		2	4 de 62
	fs-cg-gen-anteced	9	0%		2	4 de 62
	fs-cg-núm	53	2%		4;2	13 de 62
	fs-cg-num-atributo	31	1%		2	12 de 26
	fs-cg-sujeto-pred	160	7%		5;4;3;2	28 de 62
Preposición	adi-prep-de	38	2%	10%	3;2	11 de 62
	fs-prep-de	37	2%		3;2	12 de 62
	fs-prep-a	59	2%		5;3;2	12 de 62
	adi-prep-a	16	1%		3	4 de 62
	omi-prep-a	58	2%		4;3	14 de 62
	fs-prep-para	18	1%		2	5 de 62
	omi-prep-en	7	0%		2	4 de 62
	fs-prep-por	13	1%		2	4 de 62
Artículo	omi-art-def	152	6%	6%	4;3;2	28 de 62
Verbo	fs-infinitivo	7	0%	2,5%	2	4 de 62
	fs-modo-indic	15	1%		2	6 de 62
	fs-modo-sub	22	1%		3;2	8 de 62
	fs-verbo-estar	13	1%		2	5 de 62
Total		2412	100%		7 textos	62 sujetos

TABLA No.5: Errores recurrentes de toda la muestra

5. Conclusiones

Este artículo ha centrado su atención en la temática de la frecuencia y recurrencia de errores en ELE basado en un análisis de Errores asistido por el Computador. Los errores pueden clasificarse según su frecuencia y sistematicidad, permitiendo así un

análisis que contribuye en los procesos de corrección y mejoramiento de la habilidad escrita en ELE. Una contribución interesante al área de la adquisición del ELE es poder determinar con mayor precisión los errores que deberían tratarse y focalizarse con tratamientos efectivos para su corrección, como los sugeridos en el ámbito del feedback correctivo escrito. Los resultados indican que los errores más *frecuentes* están relacionados con la *falsa selección del género*, el uso incorrecto de la *acentuación de las palabras esdrújulas, agudas y diacríticas y* la *falsa selección de uso en la preposición*. En esta misma línea, los errores más *recurrentes* se registran en la *acentuación de las palabras esdrújulas y falsa selección del género gramatical* (aunque en orden inverso). Luego, en tercer lugar está la *problemática acentual en los hiatos*, que no coincide con los errores más frecuentes. Finalmente, los dos últimos errores coinciden también con los resultados obtenidos en las frecuencias generales, esto es: *la omisión de las tildes en los diacríticos*, y *la falsa selección de la preposición*.

Finalmente, consideramos que el diseño e implementación del Corpus CAELE se constituye en un aporte valioso en el ámbito de los corpus de textos escritos en formato computacional en el ELE. Esto posibilita la investigación y el análisis de producciones escritas reales de ELE y sus problemáticas lingüísticas, enriqueciendo así también el ámbito de la adquisición y enseñanza del ELE. El Corpus CAELE se está incrementando semestralmente con la recolección de nuevos textos en formato electrónico de aprendientes de ELE.

Referencias Bibliográficas

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

CAMPILLOS, L. (2014): "Las preposiciones en el habla no nativa de nivel intermedio: análisis de la interlengua basado en corpus", *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas,* 16. *https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_52f75b4406331.pdf* [consulta: 20/09/2017].

CORDER, S. P. (1975): "Error analysis, interlanguage and second language acquisition", *Language Teaching,*. 8, n. 4, 201-218.

DAGNEAUX, E.; DENNESS, S.; GRANGER, S. (1998): "Computer-aided error analysis", *System*, 26, n.2, 163-174.

ELLIS, R. (1997): *Second language Acquisition,* Oxford: Oxford University Press.

ELLIS, R., SHEEN, Y., MURAKAMI, M., y TAKASHIMA, H. (2008): "The effects of focused and unfocused written corrective feedback in an English as a FL context", *System*, 36 (3), 353-371.

FERREIRA, A. (2011). Proyecto de Investigación FONDECYT No. 1110812 "Un sistema tutorial inteligente para la focalización en la forma en ELE".

FERREIRA, A. (2014). Proyecto de Investigación FONDECYT No. 1140651. El *feedback* correctivo escrito directo e indirecto en la adquisición y aprendizaje del ELE.

FERREIRA, A. y LAFLEUR, N. (2015): "Analyse et description des erreurs prépositionnelles les plus fréquentes en espagnol L2", *Lingüística y Literatura*, 68, 57-79.

FERREIRA, A, ELEJALDE, J. y VINE. (2014): "Análisis de Errores Asistido por Computador basado en un Corpus de Aprendientes de Español como Lengua Extranjera", *Revista Signos*, 47(86), 385-411.

FERREIRA, A. (2003): , Ph.D. Thesis, University of Edinburgh, UK.

GRANGER, S. (2003): "The international corpus of learner English: a new resource for foreign language learning and teaching and second language acquisition research", *Tesol Quarterly*, 37, n.3, 538-546.

GRANGER, S. (2004): "Computer learner corpus research: current status and future prospects", *Language and Computers*, 52, n. 1, 123-145.

LOZANO, C. y MENDIKOETXEA, A. (2013): "Learner corpora and second language acquisition", A. Díaz-Negrillo, N. Ballier & P. Thompson (ed.): *Automatic treatment and analysis of learner corpus data*, Amsterdam: John Benjamins, 65-100.

O´DONNELL, M. (2008): "Demonstration of the UAM Corpus Tool for text and image annotation". Proceedings of the ACL-08: HLT Demo Session (CompanionVolume), Columbus, Ohio, June. *Association for Computational Linguistics,* 13–16. https://www.uam.es/proyectosinv/woslac/DOCUMENTS/Presentations%20and%20articles/ODonnellACL08.pdf [consulta: 20/09/2017].

PIÑOL, M. (2012): *Lingüística de corpus y enseñanza del español como 2/L,* Madrid: Arco/Libros.

QUIÑONES, A. (2009a): "Análisis de errores ortográficos en aprendices: alemanes de español como lengua extranjera", *Philologia hispalensis*, 23, 17-36.

QUIÑONES, A. (2009b): "El análisis de errores en el campo de ELE. Algunas cuestiones metodológicas", *Revista Nebrija de Lingüística aplicada a la Enseñanza de Lenguas*, *(5), https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_5316e9a63dc81.pdf* [consulta: 20/09/2017].

ROJO, G y PALACIOS, I. (2015): "Corpus de aprendices de español (CAES)", *Journal of Spanish Language Teaching,* 2, n.2, 194-200.

TORIJANO, J. A. (2006): "Lo que nos enseñan los errores", *Signum: Estudos da Linguagem,* 9, n.1, 141-206.

VÁZQUEZ, G. (2009): "Análisis de errores, el concepto de corrección y el desarrollo de la autonomía", *Revista Nebrija de Lingüística aplicada a la Enseñanza de Lenguas,* 5 . https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_5319adfdd414b.pdf [consulta: 20/09/2017].