

## ¿POR QUÉ LOS ESTUDIANTES ANGLÓFONOS COMETEN MÁS ERRORES DE ATRIBUCIÓN DE GÉNERO EN NIVELES AVANZADOS DE ESPAÑOL?<sup>1</sup>

*Why do English-speaking students make more gender  
attribution errors at advanced levels of Spanish?*

NATIVIDAD HERNÁNDEZ MUÑOZ

*Universidad de Salamanca*

*natih@usal.es*

0000-0003-2773-9545

CARMELA TOMÉ CORNEJO

*Universidad de Salamanca*

*carmela\_tome@usal.es*

0000-0002-1770-3752

### RESUMEN

El presente trabajo examina la persistencia de los errores de atribución de género en la producción escrita de aprendices anglófonos de español, a partir del análisis pseudolongitudinal de textos pertenecientes al corpus COWS-L2H. Los resultados obtenidos confirman que esta resistencia a la mejora no puede explicarse únicamente en términos de interferencia interlingüística o desconocimiento gramatical, sino que se relaciona con la consolidación léxica incompleta, la fosilización de formas incorrectas y la aplicación inestable de reglas morfológicas, especialmente en relación con sustantivos con terminaciones no prototípicas. Desde una perspectiva pedagógica, se subraya la necesidad de abordar el género como una

<sup>1</sup> Este trabajo ha sido financiado por la Unión Europea a través del proyecto MultiLingualGender (MSCA Staff Exchanges, Horizonte Europa, acuerdo de subvención N.º 101182959).

propiedad léxica idiosincrática que debe ser integrada de forma explícita y sostenida en la instrucción, incluso en niveles intermedios y avanzados.

Palabras clave: español como lengua extranjera (ELE); corpus de aprendices; errores gramaticales; género gramatical.

#### ABSTRACT

The present study examines the persistence of gender attribution errors in the written production of English-speaking learners of Spanish, based on a pseudolongitudinal analysis of texts from the COWS-L2H corpus. The results confirm that this resistance to improvement cannot be explained solely in terms of cross-linguistic interference or grammatical ignorance, but is instead related to incomplete lexical consolidation, the fossilization of incorrect forms, and the unstable application of morphological rules, particularly when dealing with nouns with non-prototypical endings. From a pedagogical perspective, the study emphasizes the need to address gender as an idiosyncratic lexical property that must be explicitly and consistently integrated into instruction, even at intermediate and advanced levels.

Keywords: Spanish as a Foreign Language (SFL); learner corpus; grammatical errors; grammatical gender.

## 1. INTRODUCCIÓN

EN UN ESTUDIO RECIENTE (Hernández Muñoz *et al.* 2026), identificamos un fenómeno llamativo al analizar la interlengua de aprendices anglófonos de español a través de los errores anotados en el corpus de aprendices COWS-L2H. Entre los aspectos estudiados, todos ellos relativos al sintagma nominal, los errores de género (concordancia y atribución) eran los que menos mejoraban con el aumento del nivel de competencia. Y, en concreto, los errores de atribución de género —es decir, aquellos en los que el estudiante asignaba incorrectamente el género gramatical a un sustantivo— no solo no disminuían de forma significativa en los cursos más avanzados, sino que, en algunos casos, tendían incluso a incrementarse. Los fenómenos de concordancia y atribución en el COWS-L2H ya habían sido estudiados previamente por Hidalgo (2025), si bien en el subcorpus de hablantes de herencia. En este último trabajo también queda patente cómo, para este tipo de hablantes, los errores de concordancia se van reduciendo lenta pero progresivamente, mientras que los de atribución persisten en mayor medida.

En su momento argumentamos que la persistencia de los errores de atribución no podía explicarse únicamente como interferencia inicial o limitación del conocimiento gramatical, sino que parecía estar relacionada con el desarrollo del léxico, con dificultades para consolidar la categoría de género gramatical en sustantivos de final no canónico y con una fosilización de formas incorrectas.

En este contexto, el objetivo de este trabajo es comprobar dichas hipótesis mediante un análisis pseudolongitudinal más detallado de los errores de atribución de género en el corpus COWS-L2H. Al hacerlo, se espera no solo aportar una explicación al aparente aumento de estas desviaciones en estudiantes avanzados, sino también ofrecer claves pedagógicas para abordar de forma más eficaz uno de los problemas gramaticales más resistentes en la enseñanza del español a anglófonos.

## 2. LOS CORPUS DE APRENDICES DE ESPAÑOL: EL COWS-L2H

Los corpus de aprendices son recopilaciones digitales de producciones orales o escritas realizadas por estudiantes de lenguas extranjeras de distintos niveles y procedencias, construidas siguiendo criterios de diseño bien definidos (Granger 2009; Granger & Gilquin & Meunier 2015; Sánchez Rufat 2015; Ferreira *et al.* 2022), cuyas interfaces de recuperación permiten acceder a los textos filtrando variables socioculturales de los aprendientes o tipologías textuales, lo que facilita su utilización por investigadores, docentes, creadores de materiales didácticos o estudiantes.

Numerosos estudios han abordado las utilidades de estos corpus tanto en contextos educativos como en la investigación en adquisición y didáctica de lenguas (Castello & Arckley & Coccetta 2015; Palacios & Barcala & Rojo 2019; Le Bruyn & Paquot 2020; Sánchez-Gutiérrez & De Cock & Tracy-Ventura 2022). Estas investigaciones describen desde análisis centrados en la interlengua y la evolución de los procesos de adquisición, así como comparaciones con el uso de hablantes nativos, hasta propuestas didácticas concretas, como la elaboración de materiales, la integración de los corpus en programas curriculares, la evaluación o la formación del profesorado.

Aunque los primeros corpus de aprendices se desarrollaron en el ámbito del inglés como segunda lengua (Granger 1998; Granger & Hung & Pech-Tyson 2002), en las últimas décadas han experimentado un gran crecimiento en el ámbito del español como lengua extranjera (Alonso Ramos 2016; Parodi & Cantos-Gómez & Howe 2022; Sánchez-Gutiérrez & De Cock & Tracy-Ventura 2022). Si bien no es objetivo de estas páginas ofrecer un inventario exhaustivo de los corpus de aprendices de español –para ello, véanse el *Indexador de Corpus de Aprendices de Español* de la Universidad Complutense de Madrid (ICAE), en Díaz Sánchez & Fernández-Pampillón & Amenós Pons (s/f), Rojo *et al.* (2022) o Hernández Muñoz *et al.* (2026)–, es posible clasificarlos, entre otros criterios, según la variedad de lenguas maternas que recogen. Así, por un lado, se encuentran aquellos corpus que incorporan producciones de aprendices con diversas lenguas de origen, como el CAES, *Corpus de aprendices de español como lengua extranjera* (Palacios *et al.* 2019); el CEDEL2, *Corpus Escrito del Español como L2* (Lozano 2022); el CORANE, *Corpus para el Análisis de Errores de Aprendices de E/LE* (Cestero & Penadés

2009), o el CAELE, *Corpus de Aprendientes de Español como Lengua Extranjera* (Ferreira & Elejalde & Blanco 2022). Por otro lado, están aquellos que se centran en aprendices de una lengua materna específica, como el CELEN, *Corpus de ELE en Japón* (Valverde 2020); el CINEAS, *Corpus de Interlengua Española de Aprendices Sinohablantes* (Calero & Serrano & Gómez Devís 2020); el CORESPI, *Corpus del español de los italianos* (Bailini 2018), o el SAELE, *Corpus de Aprendientes Suecos de ELE* (Pino Rodríguez 2012), entre otros muchos.

En este último grupo se inserta el corpus de aprendices COWS-L2H (*Corpus of Written Spanish of L2 and Heritage Speakers*), desarrollado en la Universidad de California Davis, EE. UU. (UC Davis). Se trata de una recopilación de 5383 textos con 1 367 258 palabras, escritos por casi 2953 hablantes de español como segunda lengua y hablantes de herencia, que combina datos longitudinales (evolución de un mismo estudiante a lo largo de su aprendizaje de español en los sucesivos cursos recibidos en la universidad) y pseudolongitudinales o transversales (datos de grupos de estudiantes de diferentes niveles) (Yamada *et al.* 2020, Sánchez-Gutiérrez *et al.* en prensa).

Hasta la fecha, las redacciones solicitadas se corresponden con los siguientes temas: «Una persona famosa», «Unas vacaciones perfectas», «Una persona especial en tu vida», «Una historia terrible», «Tú mismo», «Una experiencia bonita», «Chaplin» y «Un lugar que te disguste». Todos estos temas cuentan con textos anotados de acuerdo con una taxonomía de errores gramaticales previamente definida (Hernández Muñoz *et al.* 2026) y se han proporcionado versiones corregidas que, junto con las originales, se han empleado para entrenar sistemas de procesamiento del lenguaje natural para la corrección de errores gramaticales (Davidson *et al.* 2020). Estas tareas de corrección y anotación de errores han sido llevadas a cabo por un equipo de investigadores de la Universidad de Salamanca entre 2020 y 2023.

### 3. LA ANOTACIÓN DE ERRORES: LOS ERRORES DE ATRIBUCIÓN DE GÉNERO

La lingüística de corpus recurre habitualmente a procedimientos informáticos automatizados para recopilar, estructurar y anotar los textos de los corpus, como, por ejemplo, para el etiquetado de las categorías gramaticales (Van Rooy 2015). No obstante, en el tratamiento de los corpus de aprendices hay determinadas tareas que todavía no se han conseguido automatizar completamente, entre ellas, la detección y clasificación de los errores (Díaz Negrillo & Fernández Domínguez 2006; Calero & Serrano & Gómez Devís 2020). Esta situación implica la necesidad de un equipo humano que tome decisiones sobre cuáles son los errores, las taxonomías de desviaciones identificadas y los objetivos científicos del etiquetado (Lüdeling & Hirschmann 2015; Díez-Bedmar 2015, 2021). Se trata de un proceso

complejo, que requiere un amplio equipo de profesionales bilingües y que supone un elevado gasto económico y temporal, por lo que la mayor parte de los corpus de aprendices de español no incorporan anotaciones de errores y, si lo hacen, suelen ser conjuntos de textos menores y focalizados en fenómenos concretos (Rojo & Palacios & Sampedro & Marsily 2022).

El proceso de anotación de errores presenta varios debates abiertos. Uno de los más complejos es la necesidad de establecer un estándar prescriptivo como punto de partida del etiquetado, lo que plantea serias dificultades en lenguas pluricéntricas como es el español. En el caso del COWS-L2H, consideramos desviaciones las señaladas en las obras académicas de referencia (fundamentalmente, *NGLE* 2025). Otro de los problemas epistemológicos es la propia conceptualización del error entendido como desviación de la norma. Por este motivo, algunos autores prefieren hablar de formas no codificadas o no convencionales (en nuestro caso, mantendremos la terminología habitual, entendiendo que estamos ante etapas necesarias y positivas en el desarrollo de la interlengua).

En el sistema de anotación del COWS-L2H, a cada secuencia incorrecta se asocia un código de error y a continuación se proporciona su corrección. Para el diseño del plan de etiquetado se han seguido las fases y recomendaciones propuestas por Lüdeling y Hirschmann (2015) y Díez-Bedmar (2015, 2021). La descripción detallada del sistema jerárquico de etiquetado, así como su validación y fiabilidad (intra e intercorrector), puede consultarse en Hernández Muñoz *et al.* (2026). En términos generales, los errores identificados abarcan aspectos gramaticales clave en la enseñanza de español como lengua extranjera, especialmente con alumnos angloparlantes: problemas de concordancia, atribuciones incorrectas de género o número, uso indebido o ausencia de pronombres y artículos, confusión entre preposiciones y verbos copulativos, colocación errónea del adjetivo y selección incorrecta de formas verbales (por tiempo, modo, persona, número o raíz verbal).

En cuanto a los errores de atribución de género, se considera que se producen cuando el estudiante asigna género a una palabra invariable (*Me gusta mucha*), cuando atribuye un género indebido a un sustantivo (*el gran casa rojo, una problema complicada, quiero ser enfermero* [con referente femenino]), o bien modifica el tema nominal haciendo que la marca de palabra coincida con la terminación esperable en masculino (*un pequeño problema*) o femenino (*una modelo muy guapa*). En estos casos, la etiqueta seleccionada es *gat* (*gender attribution*). Como en el resto del etiquetado, la forma proporcionada por el aprendiz se coloca entre corchetes, seguida de la forma correcta (entre llaves) y del código de error, en el que se distinguen dos niveles: tipo de error (*gat*, en estos ejemplos) y la categoría gramatical afectada (noun, det, adj, etc.). Así aparecerían en el corpus algunos de los errores previos:

- (1) Me gusta [mucha]{mucho}<gat:adv> Rihanna  
 (2) Yo vivo en [el]{la}<gat:det> gran casa [rojo]{roja}<gat:adj>  
 (3) Tuvo un pequeño [problemo]{problema}<gat:noun>

#### 4. METODOLOGÍA

Para este estudio hemos seleccionado un total de 491 textos correspondientes al tema «Tú mismo» («*Yourself*»), en vista de que, según nuestros resultados previos (Hernández Muñoz *et al.* 2026), es el tema que concentra un mayor número de errores asociados al sintagma nominal. Al igual que en el trabajo citado, todos los textos pertenecen a estudiantes de español como lengua extranjera (no a hablantes de herencia) procedentes de distintos niveles de los cursos de español general de la Universidad de California Davis. En concreto, se toman 350 textos de nivel 1 (N1) –nivel Introdutorio, correspondiente a tres asignaturas de primer año y equivalente a un A1-A2– y 141 textos de nivel 3 (N3) –curso de Composición que comprende dos asignaturas de tercer año equivalentes a un B1+ / B2–. La Tabla 1 muestra los datos generales de los textos analizados y la Tabla 2, los correspondientes a los errores de atribución de género.

TABLA 1. Datos de los textos analizados bajo el tema «Tú mismo»

	N1	N3	Todos
Textos (t)	350	141	491
Palabras (p)	75996	37024	113020
Etiquetas (e)	4543	1788	6331
Promedio p*t	217	263	230,2
Promedio e*t	13,0	12,7	12,9
Promedio e*p	0,06	0,048	0,056

TABLA 2. Análisis de las etiquetas de atribución de género en valores absolutos, por texto y por palabra

	N1	N3	N1 por texto	N3 por texto	N1 por palabras*100	N3 por palabras*100
gat	137	129	0,39	0,91	0,08	0,10
gat:det	61	38	0,17	0,27	0,06	0,08
gat:adj	43	29	0,12	0,21	0,03	0,05

gat:noun	21	20	0,39	0,91	0,01	0,03
gat:pron	5	12	0,01	0,09	0,00	0,00
gat:adv	2	1	0,01	0,01	0,00	0,00
gat:in	2	1	0,01	0,01	0,18	0,35

Con el fin de delimitar el análisis y centrarnos únicamente en errores evidentes de atribución, adoptamos un enfoque más restrictivo del fenómeno. Hemos considerado únicamente los errores que afectan al género de los sustantivos (se excluyen, por tanto, los casos del tipo *Me gusta mucha*) y, hemos descartado también aquellos casos cuyo género se deriva de referentes externos (se descartan, por tanto, ejemplos del tipo *Quiero ser enfermero* [cuando es una estudiante femenina la que escribe el texto]) porque no en todos los casos disponíamos de la información necesaria para su correcta interpretación. Los errores restantes se han clasificado en tres tipos:

- Tipo 1: sustantivos con al menos dos elementos concordados en el género incorrecto:  
 Quiero trabajo en un centro de tratamiento en [la]{el}<gat:det> [próxima]{próximo} <gat:adj> año.
- Tipo 2: sustantivos masculinos en *-a* interpretados como femeninos, o bien femeninos en *-o* interpretados como masculinos:  
 Me gusta usar Canva y [otras]{otros}<gat:det> programas de diseño.
- Tipo 3: sustantivos modificados para que la marca de palabra coincida con la terminación esperable según el género atribuido:  
 Mi [equipa]{equipo}<gat:noun> [favorita]{favorito}<gat:adj> son los Boston Celtics.

## 5. RESULTADOS Y DISCUSIÓN

La Tabla 3 muestra los errores de atribución de género en sustantivos inanimados, organizados por nivel, tipo y frecuencia relativa al número de palabras. Tal y como se había adelantado, los estudiantes de nivel 3 cometen, en proporción, un mayor número de errores de atribución indebida del género gramatical (0,09 frente a 0,05), si bien este resultado se matiza al atender a su clasificación. Mientras que los errores de tipo 3, asociados a la modificación de la forma léxica, disminuyen, se produce un notable incremento en los errores de tipo 2, es decir, aquellos motivados por una interpretación errónea del género gramatical en función de la terminación del sustantivo (por ejemplo, interpretar *problema* como femenino). Esto sugiere, por una parte, cierta consolidación de las formas léxicas (las innovaciones morfológicas se reducen) y, por otra parte, una sobregeneralización de la regla gra-

matical aprendida en la L2 y no presente en la L1. A medida que los aprendientes desarrollan una conciencia morfológica, cometen más errores: la confianza en la regla aún no se acompaña de una internalización sólida de las excepciones léxicas.

Tabla 3. Análisis de los errores de atribución de género en sustantivos inanimados

	Textos	Palabras	Errores de atribución de género (valores absolutos / por palabras*100)	Errores de tipo 1 (valores absolutos / por palabras*100)	Errores de tipo 2 (valores absolutos / por palabras*100)	Errores de tipo 3 (valores absolutos / por palabras*100)
N1	350	75996	38 / 0,05	22 / 0,03	7 / 0,01	9 / 0,01
N3	141	37024	32 / 0,09	15 / 0,04	16 / 0,04	1 / 0,00

La Tabla 4 recoge los sustantivos que presentan errores en cada nivel, indicando el tipo de error y la frecuencia relativa de las ocurrencias. Se somborean las formas coincidentes en ambos niveles (*año, persona, piel, vida, día, programa, tema*). El hecho de que ciertos sustantivos sigan siendo problemáticos en el nivel más avanzado apunta a un fenómeno de fosilización léxica y a una dificultad estructural en la adquisición de género gramatical en español por parte de hablantes anglófonos.

En particular, la tendencia a aplicar estrategias de analogía formal con terminaciones prototípicas (como interpretar *programa* como femenino por su terminación en *-a*) parece mantenerse activa incluso cuando los aprendices han superado otras dificultades gramaticales.

Ahora bien, al mismo tiempo, entre los errores de tipo 1 (atribución incorrecta con al menos dos elementos concordados), el 35,5 % en el nivel 1 y el 35,7 % en el nivel 3 consisten en interpretar como masculinas palabras con marca de palabra *-a* (*casa, familia, mascota, persona, semana y vida* en el nivel 1; *biblia, característica, escuela, persona y vida* en el nivel 3). Estos errores podrían reflejar intentos de aplicar la excepción de la regla aprendida (algunos nombres en *-a* son masculinos), todavía de forma inestable. En cambio, los errores asociados al patrón contrario (masculinos en *-o* que se interpretan como femeninos) son marginales: solo un sustantivo (*año*) registra este error, si bien aparece en ambos grupos de estudiantes.

Finalmente, las modificaciones del tema recogidas en los errores de tipo 3 apuntan a una inconsistencia en la adquisición de ciertas formas léxicas, producto de cruces con formas pertenecientes a otras categorías gramaticales, como verbos (*equipo, juego, trabajo, baile*) o adjetivos (*futuro*).

Tabla 4. Sustantivos con errores de atribución de género en cada nivel

N1			N3			
Sustantivo	Tipo de error	% ocurrencias	Sustantivo	Tipo de error	% ocurrencias	
actividad	1	5,3	año	1	3,1	
año		5,3	biblia		3,1	
casa		2,6	característica		3,1	
ciudad		5,3	costumbre		3,1	
color		2,6	edad		3,1	
deporte		2,6	escuela		3,1	
elección		2,6	generación		3,1	
especialidad		2,6	hospital		3,1	
familia		2,6	lugar		3,1	
llave		2,6	mujer		3,1	
mascota		2,6	nariz		3,1	
noche		2,6	persona		3,1	
persona		5,3	piel		3,1	
piel		2,6	vida		6,3	
semana		5,3				
tradición		2,6				
vida		2,6				
día	2	7,9	clima	2		3,1
problema		2,6	día		3,1	
programa		2,6	foto		3,1	
rosa		2,6	idioma		15,6	
tema		2,6			programa	21,9
					tema	3,1
equipo	3	2,6	baile	3	3,1	
futuro		13,2				
juego		2,6				
montaña		2,6				
trabajo		2,6				

Estos resultados respaldan las hipótesis que planteamos en Hernández Muñoz *et al.* (2026) y son consistentes con el trabajo de Hidalgo (2025) con hablantes de herencia, en el que también se documenta una mayor presencia de los errores de atribución del género en sustantivos inanimados de final no canónico.

## 6. CONCLUSIÓN

Los resultados de este pequeño estudio confirman que la atribución de género gramatical representa un desafío persistente para los estudiantes anglófonos de español. Estos errores, a diferencia de los de concordancia, muestran una evolución más compleja: en niveles avanzados se mantienen errores debidos a la fosilización de ciertas formas léxicas y a las expectativas morfológicas de los aprendices. Esta situación parece indicar que el conocimiento gramatical no basta por sí solo para evitar las atribuciones indebidas; se requiere una consolidación del género como parte del conocimiento léxico específico de cada sustantivo.

Desde una perspectiva pedagógica, estos resultados subrayan la importancia de abordar el género como una característica esencial en la enseñanza del vocabulario, no solo en niveles iniciales, sino también en niveles intermedios y avanzados. El aumento del nivel de competencia no puede implicar el abandono de su presentación como una propiedad idiosincrática de cada sustantivo y su trabajo sistemático mediante estrategias de memorización léxica, reconocimiento de patrones y conciencia metalingüística sobre las irregularidades del sistema.

## REFERENCIAS BIBLIOGRÁFICAS

- ALONSO RAMOS, Margarita (2016): «Spanish learner corpus research: Achievements and challenges», en Margarita Alonso Ramos (ed.), *Spanish Learner Corpus Research: current trends and future perspectives*, Amsterdam, John Benjamins, 3-31.
- BAILINI, Sonia (2018): *Corpus del Español de los Italianos*. Recuperado de [https://corespiy-corite.altervista.org/?doing\\_wp\\_cron=1671141855.4830360412597656250000](https://corespiy-corite.altervista.org/?doing_wp_cron=1671141855.4830360412597656250000).
- CALERO FERNÁNDEZ, María Ángeles & SERRANO ZAPATA, Maribel & GÓMEZ DEVÍS, María Begoña (2020): «Codificación y etiquetado en los corpus de aprendices y su aplicación didáctica: la propuesta del Corpus de Interlegua Española de Aprendices Sinohablantes (CINEAS)», en *E-Aesla. Revista digital de lingüística aplicada* 6, 206-222.
- CASTELLO, Erik & ARCKELEY, Katherine & COCCETTA, Francesca (2015): *Studies in Learner Corpus Linguistics Research and Applications for Foreign Language Teaching and Assessment*, Bern/Berlin/Bruxelles/Frankfurt am Main/New York/Oxford/Wien, Peter Lang.
- CESTERO MANCERA, Ana María & PENADÉS MARTÍNEZ, Inmaculada (2009): *Corpus de textos escritos para el análisis de errores de aprendices de E/LE (CORANE)*, Alcalá de Henares, Universidad.
- DAVIDSON, Sam & YAMADA, Aaron & FERNÁNDEZ MIRA, Paloma & CARANDO, Agustina & SÁNCHEZ-GUTIÉRREZ, Claudia H. & SAGAE, Kenji (2020): «Developing NLP Tools with a New Corpus of Learner Spanish», en *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 7238-7243.
- DÍAZ NEGRILLO, Ana & FERNÁNDEZ DOMÍNGUEZ, Jesús (2006): «Error Tagging Systems for Learner Corpora», en *Revista española de lingüística aplicada* 19, 83-102.

- DÍAZ SÁNCHEZ, Alicia & FERNÁNDEZ-PAMPILLÓN, Ana & AMENÓS PONS, José (s/f): *Indexador de Corpus de Aprendices de Español*. Recuperado de [http://repositorios.fdi.ucm.es/corpus\\_aprendices\\_espa%c3%b1ol/view/paginas/view\\_paginas.php?id=1](http://repositorios.fdi.ucm.es/corpus_aprendices_espa%c3%b1ol/view/paginas/view_paginas.php?id=1)
- DÍEZ-BEDMAR, María Belén (2015): «Dealing with Errors in Learner Corpora to Describe, Teach and Assess EFL Writing: Focus on Article Use», en Erik Castello & Katherine Arceley & Francesca Coccetta (eds.), *Studies in Learner Corpus Linguistics Research and Applications for Foreign Language Teaching and Assessment*, Bern/Berlin/Bruxelles/Frankfurt am Main/New York/Oxford/Wien, Peter Lang, 37-70.
- DÍEZ-BEDMAR, María Belén (2021): «Error Analysis», en Nicole Tracy-Ventura & Magali Paquot (eds.), *The Routledge Handbook of Second Language Acquisition and Corpora*, New York, Routledge, 90-104.
- FERREIRA CABRERA, Anita Alejandra & ELEJALDE GÓMEZ, Jessica & BLANCO SAN MARTÍN, Lorena (2022): «Diseño e implementación del Corpus de Aprendientes de Español como Lengua Extranjera (CAELE)», en *Círculo de lingüística aplicada a la comunicación* 90, 137-155.
- GRANGER, Sylviane (ed.) (1998): *Learner English on computer*, London, Longman.
- GRANGER, Sylviane (2009): «The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation», en Karin Aijmer (ed.), *Corpora and Language Teaching*, Amsterdam, John Benjamins, 13-32.
- GRANGER, Sylviane & GILQUIN, Gaëtanelle & MEUNIER, Fanny (2015): «Introduction: Learner corpus research—past, present and future», en Sylviane Granger & Gaëtanelle Gilquin & Fanny Meunier (eds.), *The Cambridge Handbook of Learner Corpus Research*, Cambridge, University Press.
- GRANGER, Sylviane & HUNG, Joseph & PECH-TYSON, Stephanie (2002): *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching*, Amsterdam, John Benjamins.
- HERNÁNDEZ MUÑOZ, Natividad & TOMÉ CORNEJO, Carmela & BELÍO APAOLAZA, Helena S. & RECIO DIEGO, Álvaro & DAVIDSON, Sam (2026): «Anotación de errores gramaticales en el corpus de aprendices de español COWS-L2H», en *Revista española de lingüística aplicada* 39.2.
- HIDALGO GARCÍA, Miguel (2025): «Estudio del género gramatical en un corpus de aprendices de español como segunda lengua y lengua de herencia: aproximación al lenguaje inclusivo», en *Educación Multidisciplinar para la Igualdad de Género*, 249-266.
- LE BRUYN, Bert & PAQUOT, Magali (2020): *Learner Corpus Research and Second Language Acquisition: an attempt at bridging the gap*, Cambridge, University Press.
- LOZANO, Cristóbal & FERNÁNDEZ MIRA, Paloma (2022): «Designing, compiling and interrogating corpora in L2 Spanish acquisition research», en *Journal of Spanish Language Teaching* 9, 190-206.
- LÜDELING, Anke & HIRSCHMANN, Hagen (2015): «Error annotation systems», en Sylviane Granger & Gaëtanelle Gilquin & Fanny Meunier (eds.), *The Cambridge Handbook of Learner Corpus Research*, Cambridge, University Press, 135-158.
- NGLE: véase RAE & ASALE 2025.
- PALACIOS MARTÍNEZ, Ignacio & BARCALA, Francisco Mario & ROJO, Guillermo (2019): «El Corpus de Aprendices de Español (CAES) y sus aplicaciones para la enseñanza/

- aprendizaje del español como lengua extranjera», en Marta Blanco & Hella Olbertz & Victoria Vázquez Rozas (eds.), *Corpus y construcciones. Perspectivas hispánicas*, Santiago de Compostela, Universidad, 273-303.
- PARODI, Giovanni & CANTOS-GÓMEZ, Pascual & HOWE, Chad (2022): *Lingüística de corpus en español / The Routledge Handbook of Spanish Corpus Linguistics*, New York, Routledge.
- PINO RODRÍGUEZ, Ayme (2012): «El uso de combinaciones de palabras con *que* en un corpus de aprendices sucos de español como lengua extranjera», en *Anuari de Filologia. Estudis de Lingüística* 2, 211-212.
- RAE & ASALE (2025): *Nueva gramática de la lengua española*, Barcelona, Planeta: Espasa, 2.<sup>a</sup> edición (revisada y ampliada)
- ROJO, Guillermo & PALACIOS, Ignacio & SAMPEDRO, María & MARSILY, Aurélie (2022): «Los corpus de aprendices de español LE/L2: panorama actual y perspectivas futuras», en *Journal of Spanish Language Teaching* 9(2), 174-189.
- SÁNCHEZ-GUTIÉRREZ, Claudia H. & DE COCK, Barbara & TRACY-VENTURA, Nicole (2022): «Spanish corpora and their pedagogical uses: challenges and opportunities», en *Journal of Spanish Language Teaching* 9(2), 105-115.
- SÁNCHEZ-GUTIÉRREZ, Claudia & MINNILLO, Sophia & CARANDO, Agustina & DAVIDSON, Sam & FERNÁNDEZ MIRA, Paloma & SAGAE, Kenji & HERNÁNDEZ MUÑOZ, Natividad & TOMÉ CORNEJO, Carmela & RECIO DIEGO, Álvaro & BELÍO APAOLAZA, Helena (en prensa): «COWS-L2H 2.0: A longitudinal and error annotated corpus of learner and heritage Spanish», en *Second Language Research*.
- SÁNCHEZ RUFAT, Anna (2015): «Análisis contrastivo de interlengua y corpus de aprendientes: precisiones metodológicas», en *Pragmalingüística* 23, 191-210.
- VALVERDE, Pilar (2020): «Diseño y creación de un corpus de aprendices de ELE en Japón (CELEN)», en *E-Aesla* 6, 223-240.
- VAN ROOY, Bertus (2015): «Annotating learner corpora», en Sylviane Granger & Gaëtanelle Gilquin & Fanny Meunier (eds.), *The Cambridge Handbook of Learner Corpus Research*, Cambridge, University Press, 79-105.
- YAMADA, Aaron & DAVIDSON, Sam & FERNÁNDEZ MIRA, Paloma & CARANDO, Agustina & SAGAE, Kenji & SÁNCHEZ-GUTIÉRREZ, Claudia H. (2020): «COWS-L2H: A corpus of Spanish learner writing», en *Research in Corpus Linguistics* 8(1), 17-32.