

Alicia San-Mateo-Valdehita

Universidad Nacional de Educación a Distancia (UNED)

Kazuko Yonekawa

Universidad de Keio, Shonan Fujisawa Campus (SFC)

Revisión de las investigaciones clave sobre análisis de errores relacionados con el uso del artículo en la interlengua de estudiantes de español L2 hablantes de lenguas eslavas y bálticas (1991-2025)

A Review of Key Research on Error Analysis Related to the Use of the Article in the Interlanguage of Spanish L2 Students Who Speak Slavic and Baltic Languages (1991-2025)

Recibido: 11.12.2025 / **Aceptado:** 04.02.2026

Resumen: El objetivo de este estudio es revisar sistemáticamente las investigaciones sobre análisis de errores que tienen que ver con el empleo del artículo, en textos escritos por aprendices de español L2 hablantes de lenguas eslavas y bálticas (carentes de esa categoría gramatical), publicadas entre 1991 y 2025. La revisión se plantea a partir de la descripción de la metodología y la

Abstract: The aim of this study is to systematically review research on error analysis related to the use of articles in written texts produced by SL students of Spanish whose first languages are Slavic and Baltic (which lack this grammatical category), published between 1991 and 2025. The review is based on the description of the methodology and error taxonomy, as well as the corpus and

taxonomía de los errores, del corpus y del perfil de los informantes con el objetivo de (1) verificar si existe alguna tendencia común, determinar los tipos de error y los usos y contextos en los que se cometen y (2) contrastar la idoneidad de la metodología de análisis de las investigaciones. Los resultados indican que, en la mayoría de los corpus, el artículo determinado genera mayor dificultad que el indeterminado, y la omisión es el error más frecuente; no obstante, convendría realizar un análisis complementario sobre el aumento de los errores de adición a partir del nivel B. Finalmente, se confirma la necesidad de homogeneizar la metodología en futuras investigaciones para permitir comparaciones fiables.

Palabras clave: análisis de errores, artículo, interlengua, lenguas bálticas, lenguas eslavas.

1. Introducción

El objetivo de este trabajo es revisar sistemáticamente y examinar las principales investigaciones sobre análisis de errores relacionados con la categoría gramatical artículo, elaboradas entre 1991 y 2025, a partir de producciones escritas por hablantes de lenguas eslavas y bálticas que estudian español como L2 (EL2). El resultado de esta recopilación y análisis permitirá identificar, por un lado, si existe alguna tendencia en cuanto a la clase de artículo que plantea más dificultad a los aprendices cuya L1 (sea eslava o báltica) carece de artículos, cuál es el tipo de error más frecuente y en qué contexto se comete; y, por otro, ratificará o reprobará la idoneidad de la metodología de análisis.

Las lenguas eslavas y bálticas son solo algunas de las que no cuentan con esa categoría gramatical. Las orientales, como el chino, el coreano y el japonés, y las ugrofinesas, como el estonio y el finés, tampoco poseen sistema de artículos. Por su parte, el turco y el persa o farsi, aunque carecen de artículo determinado, para expresar la indefinición, emplean, respectivamente, el numeral *uno* (*bir*) —por ejemplo, *bir kitap* ‘un libro’— y el artículo indeterminado en la forma de sufijo (*-i/-yi*) añadido al sustantivo singular —*ketâb-i* ‘un libro; algún libro’—. En cambio, el único artículo que emplean las lenguas

the profile of the participants, with the objectives of (1) determining whether any common trends exist, identifying the types of errors and the uses and contexts in which they occur, and (2) assessing the suitability of the analytical methodologies employed in the studies. The results indicate that, in most corpora, the definite article poses greater difficulty than the indefinite article, and omission is the most frequent error; however, further analysis of the increase in addition errors from level B onward would be advisable. Finally, the need to standardize methodology in future research to allow for reliable comparisons is confirmed.

Keywords: article, Baltic languages, error analysis, interlanguage, Slavic languages.

semíticas es el determinado; en hebreo: $\text{—הַבַּיִת (ba-)} \text{—הַסֵּפֶר (ba-séfer)}$ ‘el libro’— y en árabe: $\text{—ال (al-)} \text{—بات كِتَاب (al-kitāb)}$ ‘el libro’—. La carencia de artículos en la L1 o de una de las dos clases podría ser la causa de la dificultad que se observa en la producción de los hablantes de EL2. No obstante, dada la diversidad de sistemas en los diferentes idiomas, en este estudio nos centraremos en los análisis de textos producidos por hablantes de las lenguas que conforman la rama baltoeslava del indoeuropeo; es decir, las eslavas y las bálticas, que comparten rasgos lingüísticos. Esta selección nos permitirá detectar comportamientos similares en los aprendices de EL2 hablantes de esa rama de lenguas y, posteriormente, se podrán comparar con los observados en textos de grupos de hablantes de otras lenguas sin artículo.

Esta investigación parte del estado de la cuestión de la tesis doctoral titulada *Análisis de errores relacionados con el uso del artículo en textos escritos por aprendientes japoneses de español como segunda lengua* (Yonekawa 2025), y amplía y actualiza la recopilación de Santos Gargallo y Alexopoulou (2021), centrada en las tesis dedicadas al análisis de errores y extraídas de las bases de datos TESEO, TDX y Dialnet; escritas en español y defendidas en universidades españolas, entre 1991 —año en el que Vázquez (1991), Fernández López (1991) y Santos Gargallo (1992) defienden la tesis— y 2019.

A partir de los criterios de la revisión de Santos Gargallo y Alexopoulou (2021), en el protocolo seguido para la compilación del corpus, por un lado, se amplía el rango cronológico hasta 2025 y, en cuanto a los tipos de investigación, además de tesis doctorales, se incluyen trabajos finales de máster (TFM) y de grado (TFG), actas de congresos y artículos, procedentes de las bases de datos Scopus (Elsevier), Web of Science (Clarivate Analytics), ERIC, Dialnet, Hispdoc y Google Académico, junto con TESEO y TDX¹ —aunque la búsqueda ha sido meticulosa, quizá algún trabajo haya quedado fuera de nuestro alcance—. Por otro lado, el análisis se restringe a los estudios sobre los errores relacionados con el empleo del artículo —o que, entre otros, tratan los de esta categoría gramatical—, en corpus de textos escritos, cuyos informantes son hablantes de lenguas eslavas o bálticas. En general, la metodología que siguen es la del Análisis de Errores o combinada con la del Análisis Contrastivo, y muestran resultados cuantitativos y cualitativos para estudiar los errores frecuentes y sistemáticos.

Con el fin de que la revisión sea rigurosa, hemos establecido de antemano los datos que se deben extraer de cada investigación: a) identificación (investigador; año; publicación: artículo, tesis...); b) corpus (disponibilidad; extensión; tema y tipo de texto: narrativo, descriptivo, epistolar...); c) informantes (L1; número; dominio de

¹ Las bases de datos ERIC (<https://eric.ed.gov/>), Dialnet (<https://dialnet.unirioja.es/>), Hispdoc (<https://hispdoc.es/>), Google Académico (<https://scholar.google.com/>), TESEO (<https://aplicaciones.ciencia.gob.es/teseo/>) y TDX (HYPERLINK "<http://www.tdx.cat/>"www.tdx.cat) son de libre acceso; en cambio, Scopus (Elsevier) y Web of Science (Clarivate Analytics) no.

la L2; perfil: edad; universitarios...); d) datos de la investigación (objeto del análisis; metodología; criterios de análisis y taxonomía; resultados cuantitativos: porcentaje de error, error más frecuente; y resultados cualitativos: contexto de error, ejemplos); conclusiones.

A continuación, explicamos la metodología y criterios del análisis de errores. Después, se destaca la importancia de la L1, se clasifican los estudios seleccionados según la L1 de los informantes y se presentan los resultados de los trabajos sobre los errores relacionados con el uso del artículo. Finalmente, exponemos las conclusiones y las líneas de investigación futura.

2. Metodología del Análisis de Errores

El objeto de estudio del método del Análisis de Errores es la desviación sistemática y frecuente debida a que el aprendiz todavía no ha incorporado algún aspecto de la L2 y, por ello, lo expresa de forma incorrecta. Además, esta línea de investigación analiza la producción real del estudiante para conseguir una visión global de su interlengua. Las ideas de Corder (1967) calaron en el ámbito del EL2 y, en los años 80 y 90, se establecieron las pautas de trabajo en las tesis doctorales de Vázquez (1991), Fernández López (1991) y Santos Gargallo (1992). Estas investigaciones se centran en examinar los errores con una actitud analítica y positiva para identificar las dificultades de los alumnos, reconstruir las diversas etapas del proceso de aprendizaje y la competencia transitoria del estudiante, evaluar el estado de interiorización de las reglas de la L2, el proceso e, incluso, los métodos de enseñanza.

El procedimiento general incluye varias etapas: a) compilación del corpus, b) identificación de los errores, c) descripción, d) clasificación, e) explicación, f) evaluación, g) discusión de los resultados y h) implicaciones didácticas para la enseñanza (Santos Gargallo 2004: 400). En primer lugar, es preciso definir el objetivo del análisis a partir de unos criterios (397): L1, tamaño de la muestra, destreza y extensión del análisis (qué aspectos se examinan: léxico, gramática, fonética...), y periodicidad de la recogida de los datos (transversal, longitudinal o pseudolongitudinal).

La técnica de recopilación de los datos del corpus está determinada por el objetivo de la investigación. En nuestro caso, como se estudian los errores detectados en la expresión escrita, se analizan corpus de redacciones sobre un tema asignado, semilibre o libre; y están disponibles si el investigador los ha incluido en su trabajo a modo de anexo. Estas investigaciones sobre errores relacionados con el empleo del artículo siguen, en primer lugar, el criterio descriptivo o lingüístico² para analizarlos y los clasifican en errores de adición y de omisión del determinante, y de selección del determinante incorrecto. En segundo lugar, en ocasiones, se aplica el criterio etiológico

² Otros criterios de análisis son el comunicativo, pragmático y cultural (Santos Gargallo 2004: 404-405).

para determinar si se trata de un error intralingual o interlingual y así explicar cómo y por qué se comete el error (por interferencia, generalización o aplicación incompleta de reglas, simplificación, evasión...). Por último, el criterio pedagógico establece, por ejemplo, si el error es transitorio, fosilizable o está ya fosilizado, o si es individual o colectivo. La dificultad para clasificar los errores estriba en que no hay consenso y cada estudio aplica la taxonomía de una forma, con lo cual la comparación entre ellos no es siempre factible.

3. Clasificación de los estudios según la L1 de los informantes

Estudiar los mecanismos de influencia de la L1 del aprendiz —y también de otras lenguas que conozca— ayuda a comprender las propiedades de la interlengua (Baralo 2004: 374-375). Aunque la transferencia de la L1 no sea algo mecánico ni todos los errores cometidos se expliquen por la diferencia entre la L1 y la L2, es una de las estrategias del aprendiz para conseguir comunicarse.

Las lenguas eslavas y bálticas constituyen la rama baltoeslava del indoeuropeo. Son dos familias con numerosas afinidades lingüísticas. Las eslavas se clasifican en tres ramas principales: eslava occidental (checo, eslovaco y polaco), eslava oriental (bielorruso, ruso y ucraniano) y eslava meridional. Las lenguas meridionales orientales son el búlgaro y el macedonio; y las meridionales occidentales, el esloveno, el croata, el bosnio, el montenegrino y el serbio (las cuatro últimas, desde el punto de vista lingüístico, se han considerado variantes del mismo idioma: el serbocroata). Las lenguas bálticas que se hablan hoy en día son las orientales: el letón y el lituano.

El hecho de que estas lenguas, menos el búlgaro y el macedonio, carezcan de artículo como categoría gramatical podría explicar la transferencia negativa y los errores relacionados con los determinantes cometidos en la L2. En ellas, la identificación del referente, que es una de las funciones del artículo determinado, se realiza mediante otros elementos lingüísticos; en serbio, con los demostrativos y pronombres (Rubinjonj 2017: 302); en ruso, con los demostrativos y posesivos (Denst 2006: 96); y, en esas dos lenguas, al igual que en polaco, también cambiando el orden de palabras y, en polaco, además, por medio de la entonación y los casos (Fernández Jódar 2006).

Dadas las características de las lenguas eslavas y bálticas, la dificultad del aprendizaje del artículo en español es un tema relevante y se han analizado los errores relacionados con el uso de esta categoría gramatical en corpus de textos escritos. Por una parte, existen trabajos en los que se estudian las producciones de grupos plurilingües que incluyen hablantes de lenguas eslavas; pero o bien el número de estos informantes no es relevante, no se ofrecen datos de sus producciones o bien se trabaja con corpus orales. De ahí que algunos estudios se hayan excluido de la revisión. Por ejemplo, Ferreira y Elejalde (2017) analizan los errores en un corpus escrito por 62 informantes de varias L1, entre ellas dos eslavas; mientras que el número de hablantes de alemán (20 sujetos;

32 %) y francés e inglés (17 de cada L1; 27 %) es relativamente similar, el de informantes checos, portugueses y suecos (2 de cada; 3 %), por un lado, y el de italianos y rusos (1 de cada; 2 %), por otro, no es representativo; y, además, no proporcionan datos ni ejemplos de las producciones de los eslavos. Por su parte, Campillos (2012) analiza los errores en un corpus oral de 40 informantes de diferentes L1, entre ellas, el polaco. Hay cuatro hablantes de cada una de estas lenguas: alemán, chino, francés, inglés, italiano, japonés, neerlandés, polaco y portugués; y un hablante de cada una de estas: coreano, finés, húngaro y turco.

Por otra parte, hay trabajos sobre producciones de grupos monolingües de lenguas eslavas o bálticas que cuentan con un nivel de dominio de EL2 desde A1 hasta C2, según el *Marco común europeo de referencia para las lenguas (MCER)* (Consejo de Europa 2002). Estos son los 15 trabajos que conforman el corpus de análisis de esta revisión, agrupados según la L1:

a) Lenguas eslavas meridionales occidentales: (1) esloveno: Santiago Alonso (2009); (2) croata: Kočman (2011) y León, Mestre *et al.* (2017); (3) serbio: Nadal (2012) y Rubinjoni (2017); (4) serbocroata: Santos Gargallo (1992).

b) Lenguas eslavas occidentales: (1) checo y eslovaco: Rodríguez García (2020); (2) eslovaco: Montoro (2013); (3) polaco: Fernández Jódar (2006) y Zimny (2016).

c) Lenguas eslavas orientales: (1) ruso: Castanedo (2019); (2) ucraniano: Lukashenko (2012).

d) Lenguas bálticas orientales: (1) letón: Maldonado (2013); (2) lituano: Rascón (2013); (3) letón y lituano: Llamas (2020).

En estas investigaciones, que cumplen los criterios establecidos para compilar el corpus (véase la Introducción), se declara seguir la metodología del Análisis de Errores o, sin indicarlo, se emplea la taxonomía de clasificación de esa metodología, de ahí que hayan sido seleccionadas. El examen de los procedimientos analíticos y de la clasificación de los errores adoptada en cada investigación evidenciará cómo se ha realizado el análisis y cuál ha sido la taxonomía de errores utilizada, para así comprobar si efectivamente la comparación de los resultados es posible.

Existen, además, otras investigaciones que quedan fuera del alcance de nuestra revisión, puesto que en ellas se analiza otro tipo de errores, como los ortográficos cometidos por hablantes de checo y eslovaco (Rodríguez García 2025); u otras cuestiones, como la de Testa (2019), sobre la relación estadísticamente significativa entre el número de artículos empleados por hablantes de polaco y el de errores cometidos, en la que, asimismo, no se especifican los tipos de error ni con qué artículo se cometen. Tampoco Belda (2004) ofrece información suficiente sobre el corpus de textos escritos por 66 universitarios rusos: no especifica el nivel de dominio de EL2, ni la extensión del corpus (número de textos y de palabras) ni el tema o tipo textual—solo indica que se registran 286 errores relacionados con el empleo del artículo, bien sea por la omisión o adición

del artículo determinado o del indeterminado, o la selección incorrecta (de este último tipo, contabiliza 58 errores por elegir el determinado en lugar del indeterminado)—. El corpus de Nikitina³ (2019) solo cuenta con siete textos de universitarios rusos y, en los resultados del análisis, no se diferencian claramente los errores cometidos por ellos de los de los informantes estonios. También se han excluido los estudios realizados con corpus orales, como el de Tarrés (2002), con hablantes de polaco, y el de De la Cavada y Sesmilo (2007), de polaco y croata; o a partir de pruebas de respuesta cerrada, como el de Martín González (1999), con informantes checos.

4. Resultados de los estudios sobre los errores relacionados con el empleo del artículo determinado e indeterminado

4.1. Estudios con informantes de lenguas eslavas meridionales occidentales

Los trabajos de Santos Gargallo (1992), Kočman (2011) y Nadal (2012) parten de corpus escritos por hablantes de lenguas eslavas meridionales occidentales; en concreto, el serbocroata, el croata y el serbio, respectivamente (véase la Tabla 1). La tesis de Santos Gargallo (1992) es la primera investigación sobre los errores léxicos y gramaticales (no solo los del empleo del artículo) detectados en la producción de hablantes eslavos. Kočman (2011) también estudia todos los errores, mientras que Nadal (2012) se centra en los relacionados con el artículo.

investigador (año)	publicación	corpus (n textos; n palabras)	L1 (n)	nivel <i>MCER</i>
Santos Gargallo (1992)	tesis	impreso (55; ~14 434)	serbocroata (n = 55)	— [B1-B2]
Kočman (2011)	artículo	no (480; —)	croata (n = 240)	A1, A2 y B1
Nadal (2012)	TFM	impreso (28; 5960)	serbio (n = 13)	B1.2

Tabla 1. Estudios con informantes de lenguas eslavas meridionales occidentales (serbocroata, croata y serbio)

Santos Gargallo (1992) y Nadal (2012) utilizan la metodología del Análisis de Errores y del Análisis Contrastivo. Por su parte, Kočman (2011) solo menciona el Análisis

³ Nikitina (2019) analiza, además, los errores relacionados con el empleo del artículo en un corpus de hablantes de estonio, al igual que Eller (2022) y Rapún (2025). El estonio no es una lengua indoeuropea, sino ugrofinesa, emparentada con el finés y el húngaro, de ahí que estos trabajos queden excluidos de la revisión.

investigador (año)	publicación	corpus (n textos; n palabras)	L1 (n)	nivel <i>MCER</i>
Eller (2022)	tesis	impreso (37; 4526)	estonio (n = 37)	B1-B2 y C1-C2
Nikitina (2019)	TFM	no (27; —)	estonio (n = 27)	A1-A2
Rapún (2025)	tesis	no (345; 65 690)	estonio (n = 15)	A1

Contrastivo, pero, al igual que Santos Gargallo (1992) y Nadal (2012), para clasificar los errores del artículo maneja la taxonomía que incluye la omisión, la adición y la selección; y, además, como Nadal (2012), contabiliza como error del artículo la falta de concordancia con el sustantivo —aunque este error no tiene que ver con las reglas de uso del determinante, sino que es una cuestión morfológica—. Santos Gargallo (1992: 235-238), en cambio, incluye ese error en otro apartado de la clasificación y no lo contabiliza como error del artículo.

Con el fin de que la comparación entre los resultados de los diferentes análisis sea posible y poder concluir qué tipo de error relacionado con el uso del artículo —es decir, error de omisión y de adición de cada clase de artículo, y error de selección entre el determinado y el indeterminado— es más frecuente, se ha aislado el número de errores que tienen que ver con la concordancia y no se ha incluido en el cómputo total para hallar el porcentaje de cada tipo de error (véanse, por ejemplo, los casos erróneos en el corpus de Nadal en la Tabla 3).

Los informantes del corpus de Santos Gargallo (1992) son 55 universitarias serbocroatas, de entre 18 y 24 años, de 3.º y 4.º de la licenciatura en Filología, de la Universidad de Zagreb (Croacia) y de la de Sarajevo (Bosnia y Herzegovina), durante el curso 1988-1989. Sobre el nivel de dominio, se indica que las estudiantes asisten a clase de español tres o cuatro horas por semana (1992: 212) y, por el curso en el que se encuentran y analizando los textos a partir de los indicadores del *Plan curricular del Instituto Cervantes (PCIC)* (Instituto Cervantes 2006), se desprende que el nivel de dominio es, en general, B1-B2, aunque no está especificado —el *MCER* se publicó años después—. El corpus se compone de 55 textos escritos (aprox. 14 434 palabras⁴) sobre un tema asignado (*¿Cuál fue el viaje más feliz de tu vida?*) y se incluye mecanografiado en el apéndice de la tesis (1992: 276-327).

En el corpus se localizan 106 errores relacionados con el uso del artículo: el 82,08 % corresponde al artículo determinado (87 casos) y el 17,92 %, al indeterminado (19) (véase la Tabla 2). El error más frecuente lo causa la adición del artículo determinado (40; 37,74 %) y lo siguen los errores de omisión de este artículo (36; 33,96 %); y, después, los de selección del determinado donde se debería utilizar el indeterminado (11; 10,38 %).

tipo de error	clase de artículo	n errores	%	%
adición	determinado	40	37,74	
omisión	determinado	36	33,96	
selección	determinado por indeterminado	11	10,38	
Casos erróneos	del artículo determinado	87		82,08
adición	indeterminado	8	7,55	
omisión	indeterminado	7	6,60	

⁴ Cálculo a partir del texto escaneado e identificado con un programa de Reconocimiento Óptico de Caracteres.

selección	indeterminado por determinado	4	3,77	
Casos erróneos	del artículo indeterminado	19		17,92
Total		106		

Tabla 2. Casos erróneos en el corpus de Santos Gargallo (1992: 218-223)

Los informantes de Kočman (2011) son 240 estudiantes croatas, de entre 15 y 18 años, matriculados en un instituto en Zagreb (Croacia), entre 2008 y 2011, de nivel A1, A2 y B1. El corpus está constituido por 480 redacciones (no se especifica la extensión, la temática ni la tipología de los textos; ni se facilita el corpus). No se ofrece información cuantitativa sobre los errores relacionados con el uso del artículo (tipo de error más frecuente, clase de artículo más problemática...), pero sí se menciona la tendencia del tipo de error que cometen los estudiantes a medida que avanzan de nivel: primero, son errores de omisión del artículo determinado; después, de selección entre determinado e indeterminado y, a continuación, de adición y de omisión del determinado.

Los informantes del corpus de Nadal (2012) son 13 estudiantes serbias (no se especifica la edad) del 4.º curso de Filología Francesa de la Universidad de Novi Sad (Serbia); en concreto, están matriculadas en el curso semestral de verano, Španski jezik VIII (Español VIII) y cuentan con nivel B1.2. El corpus se compone de 28 textos (5960 palabras) sobre un tema libre (viajes, anécdotas, eventos...), procedentes del diario o cuaderno de trabajo; y se incluye en un anexo del TFM (55-68). En este estudio, se recogen también 15 muestras de una prueba de gramática de completar huecos con los determinantes, unos cuestionarios y entrevistas personales para verificar los resultados.

En el corpus escrito, se identifican 35 errores relacionados con el uso del artículo —más tres de concordancia de género, que no se incluyen en el cómputo total de errores del empleo del artículo, pues es un error morfológico—: el 91,43 % (32 casos) tiene que ver con el uso del artículo determinado y el 8,57 % (3), con el del indeterminado. El tipo de error más frecuente se debe a la adición del determinado (18), que representa el 51,43 %; lo siguen los errores de omisión de ese mismo artículo (11; 31,43 %); y, en tercer lugar, los errores de selección del artículo determinado en lugar del indeterminado (3; 8,57 %) (véase la Tabla 3).

tipo de error	clase de artículo	n errores	%	%
adición	determinado	18	51,43	
omisión	determinado	11	31,43	
selección	determinado por indeterminado	3	8,57	
Casos erróneos	del artículo determinado	32		91,43
omisión	indeterminado	2	5,71	
adición	indeterminado	1	2,86	

Casos erróneos	del artículo indeterminado	3		8,57
Total		35		
concordancia de género	determinado e indeterminado	3		

Tabla 3. Casos erróneos en el corpus de Nadal (2012: 32)

4.2. Estudios con informantes de lenguas eslavas occidentales

Los trabajos realizados a partir de corpus escritos por hablantes de lenguas eslavas occidentales son los de Rodríguez García (2020), de checo y eslovaco; Montoro (2013), de eslovaco; y Fernández Jódar (2006) y Zimny (2016), de polaco (véase la Tabla 4).

investigador (año)	publicación	corpus (n textos; n palabras)	L1 (n)	nivel <i>MCER</i>
Rdez. G. ^a (2020)	tesis	impreso checo (192; 38 627) eslovaco (201; 45 600)	checo (n = 192) eslovaco (n = 201)	B1, B2 y C1
Montoro (2013)	tesis	escaneado (muestra) (126; ~14 238)	eslovaco (n = 42)	B1, B1+/B2 y C1
Fdez. Jódar (2006)	tesis	impreso (67; 16 267)	polaco (n = 26)	A2+, B1+ y B2+
Zimny (2016)	tesis	impreso (192; 22 498)	polaco (n = 192)	A1-A2, B1-B2 y C1-C2

Tabla 4. Estudios con informantes de lenguas eslavas occidentales (checo, eslovaco y polaco)

Zimny (2016) se centra en el análisis de errores del artículo; Rodríguez García (2020) estudia, asimismo, los demás errores morfosintácticos y los léxicos, y Montoro (2013) y Fernández Jódar (2006) analizan también los ortográficos. Siguen la metodología del Análisis de Errores y la taxonomía incluye los errores de omisión, adición y selección. Fernández Jódar (2006) y Zimny (2016) computan por separado los errores de selección de cada clase de artículo y los casos en los que se utiliza el artículo determinado en lugar del indeterminado los contabilizan como errores del primero. En cambio, Rodríguez García (2020) y Montoro (2013) no diferencian entre los errores de selección de cada artículo y, además, Montoro (2013) añade el error de selección entre el artículo y otra categoría como la preposición. Rodríguez García (2020) y Montoro (2013), entre los errores morfológicos, señalan la falta de concordancia entre el artículo y el sustantivo; en cambio, Fernández Jódar (2006) considera que es un error léxico de asignación de género. En cualquier caso, en el cómputo total de errores solo se incluyen los relacionados con el empleo de artículo (véanse los casos erróneos en el corpus de Montoro en la Tabla 6 y los del corpus de Fernández Jódar en la Tabla 8).

Los informantes del corpus de Rodríguez García (2020) son universitarios checos y

eslovacos (no se especifica la edad), matriculados en la asignatura obligatoria Jazykovy seminář I-VI (seminario sobre gramática y léxico), del grado en Lengua y Literatura Españolas, en la Universidad Masaryk de Brno (República Checa). En el curso 2014-2015 se recogen textos en el 1.º curso; en 2015-2016, en el 2.º y en 2016-2017, en el 3.º; es decir, se trata de un estudio longitudinal: los estudiantes comienzan con nivel B1 y terminan con C1.

El corpus de hablantes de checo se compone de 192 textos escritos (38 627 palabras) y el de los hablantes de eslovaco, de 201 textos (45 600 palabras), sobre temas variados (el fin de semana, las vacaciones, el mejor amigo, la película favorita, un plan de visita de un museo, una enfermedad, el sistema educativo, el profesor favorito, reseña de una novela, la religión, el futuro laboral, un buen jefe, etc.) y de diferente tipo (narrativos, descriptivos y expositivo-argumentativos). Los textos se incluyen en los anexos⁵ de la tesis (2020: 610-692).

En el corpus de hablantes de checo, se localizan 503 errores relacionados con el uso del artículo, que suponen el 30,73 % de los errores morfosintácticos (1637). El tipo de error más frecuente lo provoca la omisión del artículo determinado (210) y representa el 41,75 % de los casos erróneos relacionados con el uso del artículo; lo siguen los errores de adición del determinado (141, 28,03 %); y, en tercer lugar, los errores de selección del artículo (no se especifica el número de cada clase) (77; 15,31 %) (véase la Tabla 5)

tipo de error	clase de artículo	checo		eslovaco	
		n errores	%	n errores	%
omisión	determinado	210	41,75	254	41,23
adición	determinado	141	28,03	203	32,95
selección	determinado e indeterminado	77	15,31	95	15,42
omisión	indeterminado	55	10,93	44	7,14
adición	indeterminado	20	3,98	20	3,25
Total		503		616	

Tabla 5. Casos erróneos en los corpus de Rodríguez García (2020: 282, 296, 309, 312, 317)

En el corpus de hablantes de eslovaco, se contabilizan 616 errores relacionados con el artículo, que representan el 29,35 % de los errores morfosintácticos (2099). Igual que en el corpus de hablantes de checo, el tipo de error más frecuente lo provoca la omisión del artículo determinado (254) y supone el 41,23 % de los casos erróneos relacionados con el empleo del artículo; en segundo lugar, se sitúan los errores de

⁵ El número de palabras de los corpus se ha calculado automáticamente a partir de estos textos de los anexos.

adición del determinado (203, 32,95 %); y, en tercer lugar, los errores de selección (95; 15,42 %) (véase la Tabla 5).

Rodríguez García (2021: 390-391) calcula el porcentaje de errores de selección sobre el total de los errores del artículo en ambos corpus (1119), pero no el número de casos. En el corpus de hablantes de checo, el error de selección del artículo determinado por el indeterminado supone el 5,3 % de los casos erróneos y el de selección del indeterminado por el determinado, el 1,51 %; y en el corpus de eslovacos, el 6,7 % y el 1,78 %, respectivamente.

En cuanto a la evolución de los errores en los tres niveles de dominio, no es posible analizarla porque Rodríguez García (2021: 388-391) presenta juntos los resultados de los errores de omisión de ambos artículos y no diferencia los errores de selección por cursos.

Los informantes del corpus de Montoro (2013) son 42 universitarias eslovacas, de entre 18 y 22 años, de los dos primeros cursos de la carrera de Español, Lengua y Cultura (todas menos dos que estudian Español en combinación con Inglés), de la Facultad de Pedagogía de la Universidad Comenius de Bratislava (Eslovaquia). Están matriculadas en los cursos 2009-2010 y 2010-2011; es decir, se trata de un análisis longitudinal: las alumnas comienzan con nivel B1 y terminan con C1.

El corpus se compone de 126 textos (aprox. 14 238 palabras) producidos en 11 actividades: cuatro de la asignatura de Ejercicios de Lengua Práctica (ELP) II (18 informantes), tres de ELP III (15) y cuatro de ELP IV (9). En los anexos de la tesis se presenta una muestra de 13 textos⁶ escritos a mano (2013: 516-534). Las producciones son de temática y tipología variadas: una presentación personal, una receta, descripciones (retrato, recuerdo de un olor y el chico ideal) y un anuncio publicitario. En este estudio, se recogen también datos de 32 exámenes de ELP III, que no se han tenido en cuenta porque no son actividades de expresión escrita⁷.

En el corpus se localizan 1106 errores y el 11,66 % (129) está relacionado con el empleo del artículo. Sin tener en cuenta los dos errores de selección entre el artículo y otra categoría gramatical, los dos tipos de error más frecuentes tienen que ver con el artículo determinado: el 45,67 % se debe a la omisión (58 casos) y el 23,62 %, a la adición (30). Y, en tercer lugar, el 13,39 % se produce por la selección incorrecta del artículo (17) (véase la Tabla 6). Esta progresión es la misma en los resultados de los textos de nivel B1; en cambio, en los de B1+/B2 y C1, es más frecuente el error de omisión del artículo indeterminado que el de selección.

⁶ El número de palabras del corpus se ha calculado a partir del promedio de palabras de estos textos (113).

⁷ Si bien la tercera actividad del examen de recuperación consiste en “escribir un correo electrónico rechazando la invitación para ir a un cumpleaños, con una extensión de entre cien y ciento veinte palabras” (2013: 301).

tipo de error	clase de artículo	ELP II (B1)		ELP III (B1+/B2)		ELP IV (C1)		Total	%
		n errores	%	n errores	%	n errores	%		
omisión	determinado	34	50	15	53,57	9	29,03	58	45,67
adición	determinado	17	25	5	17,86	8	25,81	30	23,62
selección	determinado e indeterminado	10	14,71	2	7,14	5	16,13	17	13,39
omisión	indeterminado	4	5,88	5	17,86	7	22,58	16	12,6
adición	indeterminado	3	4,41	1	3,57	2	6,45	6	4,72
Total		68		28		31		127	
selección	otra categoría	1		1		0		2	

Tabla 6. Casos erróneos en el corpus de Montoro (2013: 351-359)

Con respecto a la evolución de los errores en relación con el dominio, comprobamos que se reducen a más de la mitad: de 68 en el nivel B1 a 31 en el C1; y que hay una reducción considerable de los errores de omisión y adición del artículo determinado y de los de selección (no se diferencia entre la clase de artículo), y solo los de omisión del artículo indeterminado aumentan ligeramente: de 4 en B1 a 7 en C1.

Los informantes del corpus de Fernández Jódar (2006) son 26 universitarios polacos (no se especifica la edad), del primer ciclo de Filología Hispánica de la Universidad Adam Mickiewicz de Poznań (Polonia), de dos promociones distintas, matriculados tres cursos consecutivos, entre 2000-2001 y 2004-2005. En el 1.º curso cuentan con nivel A2+; en 2.º, B1+; y en 3.º, B2+. Como son los mismos informantes, se trata de un estudio longitudinal.

El corpus se compone de 67 textos (16 267 palabras) sobre un tema cada curso: 23 redacciones del 1.º curso (*Un día que cambió tu vida*), 26 de 2.º (*Las últimas o tus mejores vacaciones*) y 18 de 3.º (*La transformación de Polonia en los últimos 25 años*). Los textos aparecen transcritos en el apéndice⁸ de la tesis (2006: 213-228).

En el corpus, si consideramos los recuentos parciales de cada tipo de error (66-85, 87, 179), hay 2100 errores: 1304 morfosintácticos, 289 léxicos y 507 ortográficos. De los 1304 errores morfosintácticos, el 18,25 % está relacionado con el empleo del artículo (238 casos) y es el error más frecuente. Teniendo en cuenta el número total de ocurrencias (correctas y erróneas) del artículo (1772), se utiliza incorrectamente en el 13,43 % de los casos y el porcentaje de error va disminuyendo según avanza el nivel de dominio (véase la Tabla 7).

⁸ El número de palabras del corpus se ha calculado automáticamente a partir de estos textos del apéndice.

curso (nivel)	casos correctos		casos erróneos		Total
	n	%	n	%	
1.º (A2+)	345	77,35	101	22,65	446
2.º (B1+)	576	88,89	72	11,11	648
3.º (B2+)	613	90,41	65	9,59	678
Total	1534	86,57	238	13,43	1772

Tabla 7. Casos correctos y erróneos por curso en el corpus de Fernández Jódar (2006: 97-98)

Sin contar los cuatro errores del artículo neutro *lo*, el 87,18 % de los errores del artículo (204 casos) está relacionado con el empleo del determinado y el 12,82 % (30) con el del indeterminado. El tipo de error más frecuente se debe a la adición del artículo determinado (112) y representa el 47,86 %; lo siguen los errores causados por la omisión de ese mismo artículo (55, 23,5 %); y, en tercer lugar, el 15,81 % (37) son errores provocados por la selección incorrecta del artículo determinado en vez del indeterminado (véase la Tabla 8).

tipo de error	clase de artículo	n errores	%	%
adición	determinado	112	47,86	
omisión	determinado	55	23,5	
selección	determinado por indeterminado	37	15,81	
Casos erróneos del artículo determinado		204		87,18
adición	indeterminado	11	4,7	
selección	indeterminado por determinado	10	4,27	
omisión	indeterminado	9	3,85	
Casos erróneos del artículo indeterminado		30		12,82
Total		234		
	neutro <i>lo</i>	4		

Tabla 8. Casos erróneos en el corpus de Fernández Jódar (2006: 99-106)

Con respecto a la evolución de los tipos de error en los textos de los tres niveles de dominio, resulta difícil analizarla porque los datos parciales de cada curso que presenta Fernández Jódar (2006: 97-104) no coinciden siempre con los totales.

Los informantes del corpus de Zimny (2016) son 192 estudiantes polacos, de entre 16 y 25 años, de cuatro instituciones públicas de Poznań (Polonia): tres centros de educación secundaria (Gimnazjum n.º XXVI y Liceum n.º XVII y n.º VI) y la Universidad Adam Mickiewicz (Facultad de Filología Hispánica), en 2016. Se distribuyen en 11 grupos y el nivel de dominio abarca de A1-A2 a C1-C2. El corpus se compone de 192

textos (22 498 palabras) —incluidos transcritos en un anexo⁹ de la tesis (2016: 391-431)— sobre un tema asignado (*Las mejores vacaciones de mi vida*). En este estudio también se recogen datos de una prueba de gramática: un texto en el que hay que seleccionar el artículo que corresponda.

Zimny (2016) presenta la tendencia en cada corpus mediante el promedio de errores con respecto al número de ocurrencias del artículo. El porcentaje de error varía entre el 36,14 %, en el corpus 3 (A1-A2), y el 9,61 %, en el 11 (C1-C2) (véase la Tabla 9). En líneas generales, los errores disminuyen según avanza el dominio de la L2. No obstante, el porcentaje de error del corpus 8 (C1-C2) es el quinto más alto (19,78 %) y, en cambio, el del 9 (A1-A2) es el tercero más bajo (16,54 %). Esas cifras se explican porque, en el corpus 8 (C1-C2) destaca el alto porcentaje del error de omisión (11,15 %) y en el corpus 9 —y también en el 2— (A1-A2), el bajo porcentaje de ese tipo de error (8,82 y 6,84 %). Al tiempo, en el corpus 9 se registra el porcentaje más bajo del error de adición (2,65 %) —y en el corpus 2, el más alto (11,23 %)— y el más alto del error de selección (5,07 %).

corpus	curso (nivel)	casos correctos (%)	casos erróneos (%)			Total (%)
			omisión	adición	selección	
1	2.º L VI (A1-A2)	77,11	15,93	4,28	2,69	22,9
2	3.º L VI (A1-A2)	81,19	6,84	11,23	0,74	18,81
3	2.º L XVII (A1-A2)	63,86	26,18	8,04	1,92	36,14
4	3.º L XVII (A1-A2)	76,85	16,33	5,12	1,69	23,14
5	3.º G (A1-A2)	68,11	19,53	9,66	2,7	31,89
6	1.º L XVII (B1-B2)	80,58	10,89	6,02	2,5	19,41
7	2.º L XVII (B1-B2)	83,43	11,65	4,27	0,65	16,57
8	3.º L XVII (C1-C2)	80,22	11,15	6,23	2,4	19,78
9	1.º U (A1-A2)	83,47	8,82	2,65	5,07	16,54
10	2.º U (B1-B2)	88,21	4,01	5,7	2,08	11,79
11	3.º U (C1-C2)	90,39	3,11	4,4	2,1	9,61

Tabla 9. Promedio de casos correctos y erróneos en el corpus de Zimny (2016: 276)

En cuanto al tipo de error más frecuente, en ocho de los once corpus, el promedio más alto es el de omisión del determinado y, después, el de adición de ese mismo artículo (véase la Tabla 10); si bien, en el corpus 2 (A1-A2), el 10 (B1-B2) y el 11 (C1-C2), el promedio de errores de adición del artículo determinado supera al de los de omisión del determinado y es el tipo más frecuente; y en el corpus 9 (A1-A2), el promedio de

⁹ El número de palabras del corpus se ha calculado automáticamente a partir de los textos del anexo.

error de selección del artículo determinado donde se debería utilizar el indeterminado (4,72 %) es mayor que el del error de adición del determinado (2,65 %).

corpus	curso (nivel)	determinado (%)	indeterminado (%)				
		omisión	adición	selección	omisión	adición	selección
1	2.º L VI (A1-A2)	13,77	4,28	2,06	2,15	0	0,63
2	3.º L VI (A1-A2)	5,04	10,86	0,74	1,8	0,37	0
3	2.º L XVII (A1-A2)	20,54	7,35	1,37	5,64	0,70	0,55
4	3.º L XVII (A1-A2)	14,16	5,12	1,69	2,17	0	0
5	3.º G (A1-A2)	16,87	8,48	2,7	2,66	1,18	0
6	1.º L XVII (B1-B2)	10,23	6,02	2,3	0,66	0	0,2
7	2.º L XVII (B1-B2)	10,17	4,27	0,42	1,49	0	0,23
8	3.º L XVII (C1-C2)	9,76	5,65	1,35	1,39	0,58	1,05
9	1.º U (A1-A2)	6,72	2,65	4,72	2,1	0	0,35
10	2.º U (B1-B2)	3,39	5,03	1,55	0,62	0,68	0,53
11	3.º U (C1-C2)	2,68	3,83	1,16	0,42	0,57	0,94

Tabla 10. Promedio de casos erróneos por tipo en el corpus de Zimny (2016: 278)

En cualquier caso, en todos los niveles, el error más frecuente está relacionado con el empleo del artículo determinado. Zimny (2016: 279-280) constata que, teniendo en cuenta los datos del corpus completo, el porcentaje de error más alto es el de omisión del artículo determinado (8,98 %) y después el de adición de esa misma clase de artículo (5,58 %).

4.3. Estudio con informantes de lenguas eslavas orientales

Los trabajos de Castanedo (2019) y Lukashenko (2012) parten de corpus escritos por hablantes de lenguas eslavas orientales, en concreto, el ruso y el ucraniano, respectivamente (véase la Tabla 11).

investigador (año)	publicación	corpus (n textos; n palabras)	L1 (n)	nivel <i>MCER</i>
Castanedo (2019)	TFM	impreso (muestra) (30; 5030)	ruso (n = 30)	B1
Lukashenko (2012)	TFM	escaneado (12; ~2426)	ucraniano (n = 12)	B1

Tabla 11. Estudios con informantes de lenguas eslavas orientales (ruso y ucraniano)

Castanedo (2019) se basa en el Análisis de Errores y el Análisis Contrastivo, y estudia los errores relacionados con el artículo (además de otros errores morfológicos, como los de las formas verbales y el uso de los verbos *ser* y *estar*). Por su parte, Lukashenko (2012) se centra en el análisis de las características del *protolenguaje* (Bickerton 1995) y analiza errores léxicos, ortográficos, pragmáticos y morfosintácticos, entre los que se

encuentran los del artículo. Ambos distinguen los errores de omisión, adición y selección de cada artículo, y de falta de concordancia de género; pero Lukashenko (2012) no contabiliza los errores por separado de manera sistemática.

Los informantes del corpus de Castanedo (2019) son 30 estudiantes rusófonos de la Universidad de Rostov del Don (Rusia), de entre 18 y 23 años: 26 del 3.º curso del grado en Filología Hispánica y cuatro del máster en Teoría de la Traducción y Comunicación Intercultural, en 2019, y cuentan con nivel B1. El corpus se compone de 30 textos (5030 palabras) sobre un tema asignado (una ciudad visitada), que son parte de una prueba (el primer bloque es una actividad sobre los tiempos verbales; el segundo, sobre *ser* y *estar* y, el tercero, sobre los artículos). En un anexo del TFM se recoge una muestra de nueve textos (2019: 63-67). Además, en una fase previa, los informantes responden un cuestionario sobre su percepción de las dificultades en la L2.

En el corpus se localizan 92 errores relacionados con el artículo y 22 tienen que ver con la asignación del género al sustantivo —este error morfológico se excluye del cálculo del total de los errores, igual que en el resto de los estudios (véase la Tabla 12)—. Centrándonos en los 70 errores del empleo del artículo, el 87,14 % está relacionado con el determinado (61 casos) y el 12,86 % con el indeterminado (9). El error más frecuente es el que se produce por la selección incorrecta del artículo determinado en lugar del indeterminado (26 casos) y representa el 37,14 %; lo sigue, bastante cerca, el error provocado por la omisión del artículo determinado (23; 32,86 %); y, en tercer lugar, el 17,14 % se debe a la adición de ese mismo artículo (12) (véase la Tabla 12).

tipo de error	clase de artículo	n errores	%	%
selección	determinado por indeterminado	26	37,14	
omisión	Determinado	23	32,86	
adición	Determinado	12	17,14	
Casos erróneos del artículo determinado		61		87,14
omisión	Indeterminado	5	7,14	
selección	indeterminado por determinado	4	5,71	
Casos erróneos del artículo indeterminado		9		12,86
Total		70		
concordancia	determinado e indeterminado	22		

Tabla 12. Casos erróneos en el corpus de Castanedo (2019: 53)

Los informantes del corpus de Lukashenko (2012) son 12 universitarias ucranianas, de entre 19 y 21 años, del 2.º curso de licenciatura de Traducción e Interpretación de la Universidad Lingüística Nacional de Kiev (Ucrania), de nivel B1. El corpus se compone de 12 textos (aprox. 2426 palabras) sobre diferentes temas (*El secreto de sus ojos: ¿Qué es la justicia?, ¿Cine o teatro? ¿Puede un arte sustituir al otro? ¿Desaparecerán ambos?, ¿Qué películas, según tu opinión, pueden causar una mala influencia en la educación*

de los niños?, visita a una ciudad, problemas relacionados con la salud o un accidente, una carta al director de un periódico...), que se incluyen en el anexo¹⁰ del TFM (2012: 134-158). Además, en el estudio se recopilan entrevistas orales. En el corpus escrito, si tenemos en cuenta las cifras parciales de cada tipo de error (2012: 115-116), hay 104 errores morfosintácticos, que suponen el 67,1 % del total (155). El artículo es la categoría que más dificultades presenta (35 casos): el 33,7 % de los errores.

4.4. Estudio con informantes de lenguas bálticas orientales

Los trabajos realizados a partir de corpus escritos por hablantes de lenguas bálticas orientales son los de Maldonado (2013), de letón; Rascón (2013), de lituano; y Llamas (2020), de letón y lituano (véase la Tabla 13).

investigador (año)	publicación	corpus (n textos; n palabras)	L1 (n)	nivel <i>MCER</i>
Maldonado (2013)	TFM	impreso (155; 25 473)	letón (n = 155)	A2 y B2
Rascón (2013)	artículo	no (71; —)	lituano (—)	A1-A2
Llamas (2020)	tesis	no letón (267; 27 817) lituano (243; 24 771)	letón (n = 89) lituano (n = 81)	A1, A2, B1, B2, C1 y C2

Tabla 13. Estudios con informantes de lenguas bálticas orientales (letón y lituano)

Maldonado (2013) se centra en el análisis de errores del uso del artículo; mientras que Rascón (2013) y Llamas (2020) estudian, asimismo, otros errores morfosintácticos, los léxicos y los ortográficos. Siguen la metodología del Análisis de Errores y la taxonomía incluye errores de omisión, adición y selección, aunque no la aplican de igual forma. Maldonado (2013) presenta en un grupo los errores de omisión y de adición de cada clase de artículo (también indica el porcentaje de cada tipo de error por curso, sin ofrecer el número de errores), y en los errores de selección considera, por una parte, la falta de concordancia entre el artículo y el sustantivo y, por otra, la selección incorrecta entre el determinado e indeterminado. En los errores de selección del artículo determinado cuenta los casos en los que se emplea el indeterminado donde se debería utilizar aquel. Rascón (2013) proporciona ejemplos de errores de omisión y adición de los artículos, y de selección del determinado por el indeterminado, sin contabilizarlos; y, en el apartado dedicado a los errores de género, añade los de concordancia entre el artículo y el sustantivo.

Por su parte, Llamas (2020) presenta los errores de adición, omisión y selección sin indicar qué artículo está implicado, y calcula cada tipo de error en función del número de casos por cada 100 palabras. Además, en los errores de selección, incluye los del artículo, pero también los del posesivo y los de las formas apocopadas y no apocopadas.

¹⁰ El número de palabras de los corpus se ha calculado a partir de los textos (escritos a mano) escaneados incluidos en el anexo. Es posible que los textos de los informantes 1 y 2 no estén completos.

Por otro lado, en el apartado dedicado al sustantivo, contabiliza la falta de concordancia del artículo.

En la medida de lo posible, en el cómputo total de los errores detectados en los corpus de estos estudios solo se contabilizan los relacionados con el uso del artículo determinado e indeterminado (véanse los casos erróneos en el corpus de Maldonado en la Tabla 14).

Los informantes del corpus de Maldonado (2013) son 155 universitarios letones, de entre 18 y 25 años, de la Facultad de Humanidades de la Universidad de Letonia —del grado en Estudios de Lenguas Modernas y Negocios, y en Filología Inglesa y Francesa, y del máster en Lenguas y Culturas Romanas— y de la Academia de Cultura de Letonia —grado en Relaciones Interculturales: Letonia-España—, en 2013. Se distribuyen en cinco grupos: 53 de 1.º curso; 56 de 2.º; 27 de 3.º; 13 de 4.º; y 6 del máster; y cuentan con nivel entre A2 y B2. El corpus se compone de 155 textos (25 473 palabras) sobre un tema asignado (la imagen de España y los españoles). Los textos están transcritos en el apéndice del TFM (2013: 58-131).

El número de errores relacionados con el artículo, si tenemos en cuenta las cifras parciales, es 778 —el 26,44 % de un total de 2943 casos—. Sin considerar los errores relacionados con la falta de concordancia ni los del neutro *lo*, se localizan 649: el 86,9 % tiene que ver con el uso del artículo determinado (564 casos) y el 13,1 %, con el del indeterminado (85). El tipo de error más frecuente lo provoca la omisión y adición del determinado (554), que representa el 85,36 %; y lo siguen los errores causados por la omisión y adición del artículo indeterminado (66), que suponen el 10,17 % (véase la Tabla 14).

tipo de error	clase de artículo	n errores	%	%
omisión y adición	determinado	554	85,36	
selección	indeterminado por determinado	10	1,54	
Casos erróneos del artículo determinado		564		86,9
omisión y adición	indeterminado	66	10,17	
selección	determinado por indeterminado	19	2,93	
Casos erróneos del artículo indeterminado		85		13,1
Total		649		
concordancia	determinado e indeterminado	104		
omisión, adición y selección	neutro <i>lo</i>	25		

Tabla 14. Casos erróneos en el corpus de Maldonado (2013: 35-48)

Maldonado (2013: 35, 38) especifica que la frecuencia media de errores de omisión y adición del artículo determinado por redacción es de 3,57 (por texto, hay 2,66 errores

de omisión y 0,92 de adición) y la frecuencia de los errores de selección del artículo determinado (casos en los que se utiliza incorrectamente el indeterminado) es de 0,06. Por otro lado, la frecuencia media de errores de adición y omisión del artículo indeterminado es de 0,43 por texto (aquí no se proporciona la cifra de cada tipo) y la frecuencia de los errores de selección (se elige incorrectamente el determinado), 0,12 por texto (2013: 44, 47). De estos datos se deduce que el tipo de error más frecuente lo provoca la omisión del artículo determinado.

En cuanto a la evolución de los errores según avanza el nivel de dominio, resulta difícil analizarla, en primer lugar, porque no se adscribe claramente cada curso a un nivel de competencia y, en segundo lugar, porque no se calcula el número de errores de cada tipo por curso ni se comparan los tipos de error por curso. Solo se contrastan los porcentajes de error de omisión y de adición de cada artículo por separado y, en el caso del artículo determinado, se constata que el error de omisión desciende considerablemente del primer curso al máster (del 73,96 al 45 %) al tiempo que el error de adición casi se duplica (del 16,67 al 35 %) (véase la Tabla 15).

tipo de error	clase de artículo	1.º	2.º	3.º	4.º	máster
		%	%	%	%	%
omisión	determinado	73,96	69,23	47,97	53,85	45
adición	determinado	16,67	18,8	30,41	26,92	35

Tabla 15. Casos erróneos de omisión y adición del artículo determinado en el corpus de Maldonado (2013: 37-38)

Los informantes del corpus de Rascón (2013) son universitarios lituanos (no se especifica el número), estudiantes de Economía, Comunicación y Derecho en la Universidad de Vilnius (Lituania), matriculados, entre 2010 y 2012, en la asignatura de libre elección Lengua Española, y cuentan con nivel A1-A2. El corpus está formado por 71 textos¹¹ (no se especifica la extensión ni se facilitan) sobre diferentes temas (los estudios universitarios, las compras o las tiendas, las estaciones del año, la lectura y la música, una autobiografía, el deporte, el tabaco, su casa, las flores y una aventura).

En el corpus, entre los errores morfosintácticos, se detectan los relacionados con el artículo causados por la omisión o la adición del determinado y del indeterminado, y por la selección del determinado donde es preciso el indeterminado. Sin embargo, no se exponen los resultados cuantitativos de cada tipo de error ni de cada clase de artículo, sino solo algunos ejemplos.

Los informantes del corpus de Llamas (2020) son 89 hablantes de letón y 81 de lituano, de entre 11 y 70 años, con nivel de A1 a C2. Por un lado, son estudiantes matriculados (entre junio y septiembre de 2018) en varias instituciones letonas:

¹¹ Rascón (2012: 94, 96) incluye más detalles sobre el corpus: son 28 textos recopilados en otoño de 2011, 23 en primavera de 2010, 9 en otoño de 2010 y 11 en primavera de 2012.

Liepājas Universitāte, Ventspils Augstskola y Latvijas Universitāte (universidades); y Latvijas Kulturas Akadēmija y Spāņu kultūras centrs Seneca (centros de idiomas) y lituanas: Klaipėdos Ažuolyno Gimnazija (centro de enseñanza secundaria), Klaipėdos Universitetas y Kauno Technologijos Universitetas (universidades) y Viva Vilnius (centro de idiomas); y, por otro, residentes en Dinamarca, España, los Estados Unidos y Luxemburgo, de los que, entre febrero y abril de 2019, se recogieron, por medios telemáticos, los textos (recibieron por correo electrónico los enunciados de las tareas y las remitieron escaneadas al investigador). En este estudio también se recopilan entrevistas para un corpus oral.

El corpus escrito de hablantes de letón se compone de 267 textos (27 817 palabras) y el de lituano de 243 (24 771 palabras) —los textos de los corpus no se facilitan—. Cada informante realiza tres tareas escritas. La primera consiste en responder varias preguntas sobre un texto (una tarjeta postal sobre las vacaciones, un ensayo sobre el valor artístico de los videojuegos, sobre la evolución de los planes de estudio y las carreras profesionales, o sobre fe, ciencia y tecnología, entre otros) y, en las otras dos tareas, hay que escribir una redacción con las pautas dadas (anuncio de una vivienda para alquilar, texto sobre un viaje realizado para una revista, artículo sobre cómo eran antes y cómo son ahora las reuniones y celebraciones familiares, carta al director de un periódico sobre la subida del precio de las entradas a los museos, reseña sobre una revista para un blog, entre otras).

En el corpus de hablantes de letón, se registran 3141 errores y el 63,9 % (2007) son morfosintácticos. El tipo de error más frecuente en todos los niveles lo provoca la omisión del artículo (no se contabilizan por separado los errores de cada clase) (véase la Tabla 16): en los textos de nivel B2, se detectan 2,31 errores por cada 100 palabras y, en los de B1, 1,82 errores. En segundo lugar, se encuentra el error de adición (si bien, en los textos de A1, la frecuencia del error de selección es la misma: 0,62 por cada 100 palabras).

	A1	A2	B1	B2	C1	C2
n palabras	4309	3198	4332	7659	3218	5101
tipo de error						
omisión	1,76	0,9	1,82	2,31	1,58	1,27
adición	0,62	0,65	0,5	0,66	0,96	0,56
selección	0,62	0,59	0,32	0,27	0,31	0,13
selección determinado e indeterminado	0,53	0,15	0,06	0,11	0,18	0,09

Tabla 16. Frecuencia del error por cada 100 palabras en el corpus de hablantes de letón de Llamas (2020: 250, 462, 467)

En cuanto a la evolución de los errores en los textos de los diferentes niveles de dominio, en los de omisión y de adición, no se observa una diferencia llamativa entre el nivel A1 y C2 (1,76 y 1,27 errores por cada 100 palabras, y 0,62 y 0,56, respectivamente); algo que es más evidente en la reducción de los errores de selección entre las clases de artículo (0,53 y 0,09).

Calculamos el número de errores a partir de los datos de la frecuencia por cada 100 palabras (Llamas 2020: 250, 462, 467), puesto que en el resto de los estudios se trabaja con el número total de errores —aunque el cálculo de los errores en relación con un número fijo de palabras es una forma de homogeneizar los resultados y aumentaría el rigor de la comparación de los resultados de las investigaciones—. Así comprobamos que el error de omisión del artículo es el más numeroso (476 casos) (véase la Tabla 17) y representa el 63,31 % de los errores relacionados con el uso de los determinantes y pronombres; en segundo lugar, se sitúa el error provocado por la adición (179; 23,83 %); y, después, el de selección (97; 12,86 %), aunque solo 49 errores se deben a la selección equivocada de la clase de artículo, pues, en ese tipo de error se contabilizan también las confusiones entre otras categorías.

tipo de error	A1	A2	B1	B2	C1	C2	Total	%
omisión	76	29	79	177	51	65	476	63,31
adición	27	21	22	51	31	29	179	23,83
selección	27	19	14	21	10	7	97	12,86
Total							752	
selección determinado e indeterminado	22,84	4,8	2,6	8,42	5,79	4,59	49	

Tabla 17. Casos erróneos en el corpus de hablantes de letón de Llamas (2020)

En el corpus de hablantes de lituano, se registran 3489 errores y el 62,05 % (2165) son morfosintácticos. El tipo de error más frecuente en todos los niveles es el causado por la omisión del artículo (véase la Tabla 18): en los textos de nivel B2, se detectan 2,44 errores por cada 100 palabras y, en los de nivel A2, 2,26 errores. El orden de la frecuencia de los errores de adición y selección varía en los textos hasta el nivel B1 (pero en los de A2 y B1 es muy similar) y, a partir del B2, aumentan los errores de adición.

En cuanto a la evolución de los errores en los textos de los diferentes niveles de dominio, se detecta un descenso considerable de la frecuencia de los errores de omisión y de selección del artículo (2,2 y 1,18 por cada 100 palabras, y 0,6 y 0,09, respectivamente); en cambio, los errores de adición aumentan ligeramente de 0,37 a 0,51 (véase la Tabla 18).

	A1	A2	B1	B2	C1	C2
n palabras	2135	4775	3329	5038	4429	5065
tipo de error						
omisión	2,2	2,26	1,98	2,44	0,97	1,18
adición	0,37	0,85	0,60	0,93	0,67	0,51
selección	1,45	0,83	0,63	0,19	0,2	0,17
selección determinado e indeterminado	0,6	0,25	0,33	0,09	0,15	0,09

Tabla 18. Frecuencia del error por cada 100 palabras en el corpus de hablantes de lituano de Llamas (2020: 250, 276, 281)

Calculamos el número de errores a partir de los datos de la frecuencia por cada 100 palabras (Llamas 2020: 250, 276, 281) y deducimos que el error de omisión del artículo es el más frecuente (446 casos) (véase la Tabla 19) y supone el 60,67 % de los errores; lo siguen los errores causados por la adición (171; 23,21 %); y, por último, por la selección incorrecta (119; 16,12 %), no obstante, solo 52 errores se deben a la selección equivocada de la clase de artículo.

tipo de error	A1	A2	B1	B2	C1	C2	Total	%
omisión	47	108	66	123	43	60	446	60,67
adición	8	41	20	47	30	26	171	23,21
selección	31	40	21	10	9	9	119	16,12
Total							736	
selección determinado e indeterminado	12,81	11,94	10,99	4,53	6,64	4,56	52	

Tabla 19. Casos erróneos en el corpus de hablantes de lituano de Llamas (2020)

5. Resultados de los estudios sobre los errores relacionados con el empleo del artículo determinado con hablantes de lenguas eslavas meridionales occidentales

La investigación de León, Mestre *et al.* (2017), con informantes croatas; la de Rubinjoni (2017), con serbios¹²; y la de Santiago Alonso (2009), con eslovenos, se centran en los errores relacionados con el empleo del artículo determinado (véase la Tabla 20), por lo que entendemos que parten de la hipótesis de la dificultad que entraña su uso.

¹² Véase también la tesis de Zečević (2016), escrita en serbio.

investigador (año)	publicación	corpus (n textos; n palabras)	L1 (n)	nivel <i>MCER</i>
León, Mestre <i>et al.</i> (2017)	artículo	no (20 escritos; ~2000-3000; 10 orales; 5-10 minutos)	croata (—)	A1 y A2
Rubinjoni (2017)	tesis	impreso c_2011 (102; 18 665) c_2014 (94; 15 572)	serbio c_2011 (n = 102) c_2014 (n = 94)	A1-A2, B1, B2 y C1
Santiago Alonso (2009)	artículo	no (110; —)	esloveno (—)	— [B1, B1+, B2, B2+ y C1]

Tabla 20. Estudios sobre el artículo determinado con informantes de lenguas eslavas meridionales occidentales (croata, serbio y esloveno)

En los tres estudios se utiliza la metodología del Análisis de Errores y el criterio descriptivo. La taxonomía incluye errores de omisión y adición del artículo determinado, y de selección de otro determinante (en lugar del artículo). Por su parte, Santiago Alonso (2009) añade el error de concordancia de género y número, que se ha excluido del cómputo total de errores, puesto que es un error morfológico, que no guarda relación con el dominio de los usos del artículo (véanse los casos erróneos en el corpus de Santiago Alonso en la Tabla 24).

Los informantes del corpus de León, Mestre *et al.* (2017) son estudiantes croatas de español, gallego y catalán (no se especifica el número ni la edad) en el departamento de Estudios Iberorromanos de la Universidad de Zadar (Croacia) (no se indica el año), de nivel A1 y A2. El corpus se compone de 20 textos escritos, de entre 100 y 150 palabras cada uno (aprox. 2000-3000 palabras), que no están disponibles —diez de nivel A1 y diez de A2—, y diez transcripciones orales de producciones de 1 a 5 minutos —cinco de A1 y cinco de A2—, de cada una de las tres lenguas. La tarea escrita consiste en la descripción de un retrato familiar o un texto sobre un tema elegido al azar, desde la descripción de una rutina hasta una presentación. En la exposición de los resultados del análisis de errores no siempre se facilitan los datos referidos al corpus de cada lengua ni de cada tipo de texto (oral y escrito). En el corpus de los 30 textos (escritos y orales) en español, se detectan 104 errores del artículo determinado: el 61,54 % se debe a la omisión del artículo (64 casos); el 32,69 %, a la adición (34); y el 5,77 %, a la selección (6) (véase la Tabla 21). La frecuencia de los errores según el nivel de dominio y el tipo de texto solo se calcula con el corpus completo de las tres lenguas (90 textos) y no se ofrecen, por separado, los datos del análisis de los textos de los estudiantes de español. Los hablantes de nivel A1 cometen 200 errores y los de nivel A2, 171. En el nivel A1, la omisión del artículo (147 errores) es más frecuente que la adición (45), mientras que, en el nivel A2, se iguala el número de errores de omisión (78) y de adición (77). Los errores de

selección son menos numerosos: 8 en A1 y 16 en A2. Por otro lado, en el corpus escrito, se cometen 4,23 errores por texto y, en el corpus oral, 3,9; es decir, el error no varía según el tipo de producción.

tipo de error	n errores	%
omisión	64	61,54
adición	34	32,69
selección	6	5,77
Total	104	

Tabla 21. Casos erróneos del artículo determinado en el corpus de León, Mestre *et al.* (2017: 152)

Rubinjoni (2017) trabaja con dos corpus (uno, compilado en 2011, y otro, en 2014) de textos escritos por estudiantes serbios de Lengua Española y Literaturas Hispánicas, de la Universidad de Kragujevac (Serbia), de entre 20 y 29 años (corpus 2011) y entre 20 y 26 años (corpus 2014). El nivel de dominio del 1.º curso de la carrera equivale a A1-A2; el del 2.º, a B1; el del 3.º, a B2; y el del 4.º, a C1.

El corpus 2011 está compuesto de 102 redacciones (18 665 palabras) sobre un tema dado (*Un viaje soñado*), escritas por 102 informantes: 23 de 1.º curso, 20 de 2.º, 28 de 3.º y 31 de 4.º. El corpus 2014 cuenta con 94 textos (15 572 palabras) sobre un tema académico (*Comenta, de manera razonada, qué asignaturas y qué contenidos de la licenciatura te resultan más interesantes y cuáles son los que no te interesan*), escritos por 94 estudiantes: 28 de 1.º curso, 14 de 2.º, 29 de 3.º y 23 de 4.º. Los textos se incluyen transcritos en los anexos de la tesis (2017: 356-397; 404-444).

En el corpus 2011, en todos los niveles, el error más frecuente se debe a la omisión del artículo determinado —entre el 38 % en el nivel A1-A2 y el 20 % en B2—; el segundo, a la adición —entre el 9 (B1) y el 5 % (A1-A2)—; y el tercero, a la selección incorrecta del artículo —entre el 6 (B1) y el 4 % (A1-A2 y C1)— (véase la Tabla 22).

corpus	curso (nivel)	casos correctos (%)	casos erróneos					
			tipo de error	%	tipo de error	%	tipo de error	%
2011	1.º (A1-A2)	53	omisión	38	adición	5	selección	4
	2.º (B1)	65	omisión	21	adición	9	selección	6
	3.º (B2)	67	omisión	20	adición	8	selección	5
	4.º (C1)	66	omisión	22	adición	8	selección	4

Tabla 22. Casos erróneos del artículo determinado en el corpus 2011 de Rubinjoni (2017: 265)

En el corpus 2014, en los tres primeros niveles, el orden de frecuencia de los tipos de error coincide con el del corpus 2011 (si bien la diferencia entre los errores de omisión y de adición es mucho menor): el error más recurrente lo causa la omisión del artículo determinado —entre el 20 % en el nivel B2 y el 18 % en B1—, seguido del error de adición —entre el 18 % (A1-A2) y el 14 % (B2)— y de selección —entre el 4 % (A1-A2 y B2) y el 2 % (B1)—. En el nivel C1, es más frecuente el error de adición (19 %) que el de omisión (15 %); e, igual que en los demás niveles, el error de selección ocupa el tercer lugar (3 %) (véase la Tabla 23).

corpus	curso (nivel)	casos correctos (%)	casos erróneos		tipo de error		tipo de error	
			tipo de error	%	tipo de error	%	tipo de error	%
2014	1.º (A1-A2)	59	omisión	19	adición	18	selección	4
	2.º (B1)	63	omisión	18	adición	17	selección	2
	3.º (B2)	61	omisión	20	adición	14	selección	4
	4.º (C1)	63	adición	19	omisión	15	selección	3

Tabla 23. Casos erróneos del artículo determinado en el corpus 2014 de Rubinjoni (2017: 265)

Los informantes del corpus de Santiago Alonso (2009) son estudiantes eslovenos, de entre 19 y 23 años, matriculados en el curso 2008-2009 en la Universidad de Liubliana (Eslovenia). El nivel de dominio de EL2 al inicio de los cursos, según nos ha comunicado la autora, es B1 (1.º), B1+ (2.º), B2 (3.º) y B2+ (4.º); y, en el último año, C1. El corpus se compone de 110 redacciones de tipología variada (no se especifica la extensión ni se facilitan los textos): 49 textos argumentativos, 18 narrativos, 17 descriptivos, 23 informes o resúmenes y 3 TFG19F.

Sin tener en cuenta los errores debidos a la asignación errónea del género del sustantivo, en el corpus, se identifican 387 errores relacionados con el uso del artículo determinado. El tipo de error más frecuente se produce debido a la omisión del artículo (295) y representa el 76,23 % de los errores; lo sigue el error causado por la adición innecesaria (50; 12,92 %); y, en tercer lugar, el error de selección (42; 10,85 %) (véase la Tabla 24). Esta progresión se observa en casi todos los niveles; sin embargo, en el nivel B1, se contabiliza el mismo número de errores de adición y de selección; y, en el B1+, hay dos errores de adición menos que de selección.

tipo de error	1.º (B1)	2.º (B1+)	3.º (B2)	4.º (B2+)	5.º (C1)	Total	%		%		%	
	n errores	%	n errores	%	n errores	%	n errores	%	n errores	%	n errores	%
omisión	101	77,1	78	82,98	51	69,86	32	66,67	33	80,49	295	76,23
adición	15	11,45	7	7,45	12	16,44	9	18,75	7	17,07	50	12,92
selección	15	11,45	9	9,57	10	13,7	7	14,58	1	2,44	42	10,85

Total	131		94		73		48		41		387
concordancia	17		5		2		0		0		24

Tabla 24. Casos erróneos del artículo determinado en el corpus de Santiago Alonso (2009: 39)

En cuanto a la evolución de los errores en los cinco niveles de dominio, en todos ellos, el tipo de error más frecuente es el de omisión y el número de errores de los tres tipos disminuye desde los textos de nivel B1+ hasta C1+ (véase la Tabla 24).

6. Usos y contextos de los errores relacionados con el empleo del artículo

Los estudios de la revisión incluyen, en su mayoría, datos cualitativos; es decir, ejemplifican los tipos de error con fragmentos de los corpus. El origen de la complejidad del dominio del artículo son los múltiples usos que presenta esta categoría gramatical, no siempre asociados a una determinada función sintáctica. De ahí que sería esperable que, en estos estudios de corpus, los ejemplos de los diferentes tipos de error estuvieran relacionados con el uso que desempeña el artículo en cada caso, pues, en definitiva, eso explica por qué la secuencia es errónea. Así, por ejemplo, Rubinjoní (2017), además de incluir referencias a los textos de los corpus en la presentación del cómputo de cada tipo de error, facilita, en los anexos, la codificación completa de todos los tipos de error clasificados según el uso del artículo (por ejemplo, anafórico directo o asociativo, genérico, con modificadores restrictivos, etc.). En el resto de los trabajos no encontramos exposiciones tan exhaustivas ni formuladas con esa precisión, pero tanto Kočman (2011), Fernández Jódar (2006), Llamas (2020), Maldonado (2013) como Rodríguez García (2020) relacionan los ejemplos de los corpus con algunos de los valores del artículo o con la clase de palabra a la que acompaña.

Por su parte, Rascón (2013) y Santiago Alonso (2009), además del uso del artículo, emplean la función sintáctica que desempeña el sintagma nominal en el que está inserto el artículo (por ejemplo, sujeto del verbo *gustar*, complemento directo, complemento circunstancial, etc.) para describir los casos erróneos.

Sin embargo, Castanedo (2019), Lukashenko (2012), Montoro (2013), Nadal (2012) y Santos Gargallo (1992) ejemplifican cada tipo de error (por ejemplo, el de omisión del artículo determinado) sin mencionar el uso o valor del artículo; y, por último, León, Mestre *et al.* (2017) no presentan ejemplos del corpus analizado, al igual que Zimny (2016), que se centra en el análisis cuantitativo, aunque indica los usos más problemáticos y los que presentan mayor número de aciertos.

En los estudios realizados con corpus de hablantes de lenguas eslavas meridionales occidentales no se observa ninguna tendencia concreta en los errores asociados a los valores de los artículos. Rubinjoní (2017: 445) extrae, por ejemplo, el caso de (1)

del corpus 2014 de hablantes de serbio, para ilustrar el error de omisión del artículo determinado antes de la segunda mención del sustantivo, con valor de anáfora directa.

(1) Por supuesto, la gramática es parte mas importante para lengua. Para mí, (la) gramatica no es interesanta, pero es importa para hablar, entender, practicar escribir¹³.

Kočman (2011: 4) también alude en varias ocasiones al desconocimiento de los hablantes de croata de las reglas del uso del artículo, aunque no presenta sistemáticamente los errores asociados al valor del artículo; por ejemplo, dice: “Es sabido que *hay* exige el artículo indeterminado”, que explica el error de (2); y también menciona la importancia de distinguir entre algo determinado, concreto, específico y general, con el fin de utilizar correctamente el artículo determinado para generalizar y no cometer errores como en (3).

(2) En la habitación hay *el armario, *la cama...

(3) (Los) Delfines viven en el mar.

Santiago Alonso (2009), en cambio, asocia, por ejemplo, el error de omisión del artículo determinado de (4) a la función de sujeto o tópico del sintagma nominal *el protagonista*.

(4) (El) Protagonista desea vivir una nueva vida sin el agresivo de su marido
Por su parte, Nadal (2012: 33) y Santos Gargallo (1992: 223) registran, respectivamente, en los corpus de hablantes de serbio y serbocroata casos que ilustran, por ejemplo, la omisión del artículo indeterminado (5) y la selección incorrecta del artículo indeterminado por el determinado (6), sin aludir al valor del artículo.

(5) Ahí pasamos tres semanas en un camping, en (un) estadio de fútbol.

(6) [...] eso era *una experiencia que más me entusiasmó.

En los trabajos elaborados con producciones de hablantes de lenguas eslavas occidentales tampoco es posible determinar comportamientos específicos. Fernández Jódar (2006: 99-100) explica algún uso del artículo al hilo de la exposición de los errores cometidos por los hablantes de polaco. En (7) se recoge un error de adición del artículo determinado con un sustantivo que “no está actualizado ni determinado por el contexto”.

(7) [...] de vez en cuando huía hacia atrás, a lo mejor buscando *las aventuras. Rodríguez García (2020: 284-285) menciona el elemento con el que se combina erróneamente el artículo (o ante el que se omite) en los corpus de hablantes de checo y de eslovaco, por ejemplo, en (8), el indeterminado antecede a sustantivos referidos a profesiones.

(8) a. Pues, son *una maestra del nivel superior en una [escuela de] primaria (checo).

b. [...] se hizo *un piloto militar (eslovaco).

¹³ Al igual que en los trabajos revisados, los ejemplos se reflejan tal y como aparecen en los corpus.

Montoro (2013: 205), por su parte, recoge segmentos del corpus de hablantes de eslovaco que ilustran los errores cometidos, como, por ejemplo, el de omisión del artículo determinado en (9), sin especificar el valor del artículo.

(9) Se añade la carne de (la) berenjena, (los) champiñones, (el) perejil y el pan. En los trabajos a partir de textos de hablantes de lenguas eslavas orientales, tanto Castanedo (2019: 49) como Lukashenko (2012: 65) recopilan ejemplos de los corpus de hablantes de ruso y ucraniano, respectivamente, como, en (10), de adición del artículo determinado y, en (11), de omisión del indeterminado, sin aludir al uso del artículo.

(10) Viaje en *el avión con mis compañeras de Universidad

(11) La película “El secreto de sus ojos” es (una) obra de arte.

En las investigaciones realizadas con corpus de hablantes de lenguas bálticas orientales no destaca ningún error asociado a un valor concreto del artículo. Llamas (2020: 653, 688) presenta los ejemplos agrupados según el elemento al que el artículo acompaña (o ante el que se omite) en los corpus de hablantes de letón y de lituano, como en (12), donde el determinado precede innecesariamente a un topónimo.

(12) a. Quiere ver el país con las excursiones y probar comida típica *del México (lituano).

b. En Málaga y también he visitado *el Gibraltar (letón).

Maldonado (2013: 36) alude también al componente del sintagma nominal, como el sustantivo referido a toda la categoría que precisa del artículo determinado en (13), ejemplo extraído del corpus de hablantes de letón.

(13) (Los) Españoles son muy activos y alegres.

Por su parte, Rascón (2013: 97) ilustra los ejemplos del corpus de hablantes de lituano con la referencia a la función sintáctica del sintagma nominal acompañada de una indicación, como en (14a), “En complemento directo, con sustantivos no actualizados previamente”; indica simplemente la función, como en (14b), “con atributo”; o no añade más que el tipo de error, como en (14c), “omisión de indeterminado”.

(14) a. [...] he visto *las películas.

b. Cuando era pequeña quería ser *una médica.

c. [...] es (un) deporte muy interesante.

En estos estudios, el análisis cuantitativo cuyo fin es determinar cuál es el tipo de error más habitual ocupa un lugar preferente y de ahí que nuestro primer objetivo esté centrado en ese análisis. Tampoco hay que olvidar que la metodología del Análisis de Errores parte del recuento de casos erróneos.

Partiendo de la recopilación de los ejemplos procedentes de los estudios de esta revisión y aunque no todos los investigadores hayan establecido la relación entre los usos del artículo y los tipos de error localizados en los corpus de hablantes de las lenguas

eslavas y bálticas, nosotros la hemos constatado; si bien, como se ha indicado, hasta el momento no ha sido posible establecer ninguna tendencia específica de cada grupo de aprendices de EL2. A continuación, recogemos casos erróneos del artículo asociados a determinados valores que destacan en los análisis por su alta incidencia.

En primer lugar, el error de omisión del artículo determinado y el de selección del indeterminado se produce, por ejemplo, cuando el artículo posee uso anafórico o anafórico asociativo (15-16) o expresa posesión inalienable (17); uso no anafórico basado en el conocimiento general con un referente concreto (condición de unicidad) (18-20); uso endofórico con sustantivos identificados por su complemento (21-22); o uso genérico con verbos de afección y psicológicos transitivos, como *odiar* o *soportar*, e intransitivos, como *encantar* o *gustar* (23-24).

(15) Juan vive en casa. (La) Casa es grande (croata; Kočman 2011: 4).

(16) Cuando esté guisado añade el orégano, la albahaca y (el) arroz (eslovaco; Montoro 2015: 205).

(17) Es necesario lavar (las) manos todo el rato (ucraniano; Lukashenko 2012: 87).

(18) Tomamos mucho (el) sol (ruso; Castanedo 2019: 49).

(19) Una cosa que fue muy interesante fue (el) metro (serbocroata; Santos Gargallo 1992: 220).

(20) Tiene 20 años y estudia en *una universidad (checo; Rodríguez García 2020: 321).

(21) Aquel hombre era el hijo de (el) autor del cuento sin título (esloveno; Santiago Alonso 2009: 41).

(22) Era *una noche más horrible en mi vida (polaco; Fernández Jódar 2006: 105).

(23) No soporto (los) deportes de invierno [...]. Me encanta (el) calor (checo; Rodríguez García 2020: 303).

(24) Me gusta *un helado de fresa (croata; Kočman 2011: 4).

En segundo lugar, el artículo determinado se añade innecesariamente para referirse a algo específico (25-27) que, en realidad, no lo es y antes de ciertos topónimos (28), entre otros contextos.

(25) Aquí en Serbia por la mañana comemos *el jamón, *el salchichón... (serbio; Nadal 2012: 33).

(26) Sé *el inglés (croata; Kočman 2011: 4).

(27) [...] quiero leer *los libros y escuchar *la música en mi terraza (serbio; Rubinjoni 2017: 242).

(28) [...], pero yo pienso que *la Letonia es [...] (letón; Maldonado 2013: 36). Por último, la selección del artículo determinado y la omisión del indeterminado son

incorrectas cuando el sustantivo va acompañado de modificadores calificativos que lo precisan (29-30) o introduce un referente nuevo en el discurso (31-33).

(29) Venecia es (una) ciudad muy cara (serbocroata; Santos Gargallo 1992: 222).

(30) Moscú es *la ciudad encantadora (ruso; Castanedo 2019: 49).

(31) [...] hay (una) mesita de noche. [...] hay (un) armario (checo; Rodríguez García 2020: 312).

(32) Estoy buscando *el vestido (lituano; Rascón 2013: 97).

(33) En la habitación hay *el armario, *la cama... (croata; Kočman 2011: 4).

Y el error de adición del artículo indeterminado se localiza con sustantivos que denotan profesiones (34) y, sobre todo, en combinación con el adjetivo *otro* (35).

(34) Cuando era pequeña, quería ser *una médica (lituano; Rascón 2013: 97).

(35) Me gustaría organizar *una otra fiesta (letón; Llamas 2020: 689).

Somos conscientes de que esta recopilación de errores asociados a los valores del artículo no da suficiente cuenta de la complejidad que entraña el dominio de la categoría gramatical. No obstante, puede servir como referencia para futuros estudios de corpus, ya que, como se ha señalado, no siempre se incluye el análisis cualitativo ni se tienen presentes de manera sistemática los usos del artículo.

7. Conclusiones

De esta revisión de las investigaciones sobre análisis de errores relacionados con el empleo del artículo por parte de hablantes de lenguas eslavas y bálticas, en primer lugar, se desprende que, independientemente de la L1 y del nivel de dominio, el artículo determinado es más problemático que el indeterminado, incluso León, Mestre *et al.* (2017), Rubinjoni (2017) y Santiago Alonso (2009) se centran solo en esa clase en sus estudios de corpus de croatas, serbios y eslovenos, respectivamente —pero hay trabajos que no ofrecen ese dato cuantitativo, como Kočman (2011), Lukashenko (2012) y Rascón (2013)—.

En segundo lugar, por un lado, en siete investigaciones con corpus de aprendices de nivel de A1 a C2 y hablantes de lenguas eslavas meridionales occidentales (croata, esloveno y serbio), occidentales (checo, eslovaco y polaco) y bálticas orientales (letón), el error más frecuente se debe a la omisión del artículo determinado —Llamas (2020) no diferencia entre la omisión de las dos clases de artículo en los corpus de letones y lituanos—. Y, por otro lado, en tres estudios con hablantes de serbio y serbocroata, y de polaco, de nivel de A2 a B2+, es más frecuente el error de adición del artículo determinado —incremento que se observa igualmente en el nivel C1 del corpus 2014 (serbio) de Rubinjoni (2017) y en dos de los corpus de Zimny (2016) (polaco) de B1-B2 y C1-C2, aunque también en otro de A1-A2—. En tercer lugar, en el corpus de hablantes de ruso (Castanedo 2019), el error más frecuente lo causa la selección incorrecta del artículo determinado.

La posible tendencia del aumento del error de adición, probablemente por ultracorrección, a medida que avanza el nivel de dominio y la aparición del error de selección merecerían ser estudiadas con más detenimiento, así como también convendría analizar algún corpus de hablantes de otras lenguas eslavas sin artículo como el bielorruso. Además, como futura línea de investigación, proponemos la ampliación de la revisión sistemática a estudios sobre corpus de producciones de hablantes de otras lenguas sin artículo, como, por ejemplo, las orientales, para así poder comparar los resultados y determinar la existencia de tendencias comunes en los aprendices de EL2 que comparten una L1 carente de la misma categoría gramatical.

Más allá de las conclusiones cuantitativas y cualitativas, para que la comparación entre los resultados de los análisis sea fiable, es necesario aplicar una metodología rigurosa y una taxonomía de errores uniforme. Las conclusiones serían más consistentes si en los estudios se expusiera la clasificación de los errores que se va a aplicar y, en concreto, en el caso del artículo, para obtener un análisis completo, hay que considerar por separado el error de adición y de omisión del artículo determinado y del indeterminado. Esta distinción no siempre se aplica en la taxonomía (véase Llamas 2020). Por ejemplo, en el estudio de Valverde y Ohtani (2014: 237) sobre la fiabilidad de la codificación de los errores relacionados con el artículo en un corpus de aprendices japoneses, las etiquetas del esquema de anotación de los errores empleadas son: *missing (definite)*, *missing (indefinite)*, *extraneous*, *confusion*, *article is correct* y *difficult to judge*; es decir, se codifica el error de omisión del artículo determinado y del indeterminado, pero no se diferencia el error de adición (*extraneous*) y el de selección (*confusion*) de cada clase de artículo. En cambio, Rapún (2025: 105-106) sí contempla los errores de cada clase de artículo.

Por otro lado, el hecho de contar como error de uso del artículo la falta de concordancia con el género del sustantivo (una cuestión morfológica) o la vacilación que existe al asignar a una clase de artículo el error de selección (empleo del determinado por el indeterminado o del indeterminado por el determinado) provoca que los resultados se desvirtúen y que la comparación entre las investigaciones no sea clara.

En esta revisión sistemática, para corregir esa inconsistencia en la medida de lo posible y con el fin de que sea comparable el porcentaje de los errores de omisión y de adición del artículo determinado y del indeterminado, y de los errores de selección de la clase de artículo —es decir, los errores relacionados con el manejo de los usos de esta categoría gramatical—, no se incluyen en el cómputo los errores morfológicos o los relacionados con el artículo neutro *lo*. De esta manera, resultan más rigurosas las conclusiones sobre la frecuencia del tipo de error que cometen los aprendices de EL2 hablantes de lenguas eslavas y bálticas. En cuanto a la aplicación de la taxonomía del error de selección, aunque la homogeneidad en el cómputo de cada error de

selección sería más coherente, la conclusión es que, en las investigaciones examinadas, este no es el tipo de error más frecuente, independientemente del artículo que se haya seleccionado y cómo se haya contabilizado (si es que se ha diferenciado entre las dos clases de artículo).

La frecuencia del error se debería calcular basándose en algún parámetro que homogeneizara los resultados de los estudios, por ejemplo, un número de palabras fijo, como proponen Llamas (2020) y Fernández Jódar (2006), aunque no determinan la extensión necesaria del corpus para que las conclusiones sean fiables. En otros trabajos de este ámbito (Thewissen 2020), se opta por calcular el porcentaje de errores de un tipo en relación con el total de errores del corpus o con el total de palabras del corpus, aunque, según Thewissen (2013: 81), “it is doubtful whether total tokens constitute the most suitable denominators for all error types”; y la investigadora aplica el método que denomina *potential occasion analysis*. De manera que, por ejemplo, los errores del artículo se contabilizarían teniendo presente el número de ocurrencias (correctas y erróneas) de esa categoría gramatical, puesto que son las ocasiones potenciales de cometer el error. Hawkins y Buttery (2010), en su análisis de errores cometidos por aprendices de inglés, primero calculan los errores por cada 100 000 palabras y después, de forma más fiable, cuantifican cada tipo respecto a los elementos de cada categoría y así consiguen una idea más clara de la proporción de aciertos y de errores de cada tipo. Apuntan que el número de ocurrencias de cada categoría puede variar considerablemente en cada 100 000 palabras y lo mismo podría suceder con los errores relacionados con el empleo de esa categoría. Sin embargo, con el segundo cálculo, “we achieve a better reflection of the universe of relevant instances within which we can quantify success or failure” (2010: 15). Siguiendo esta propuesta, Rapún analiza los errores relacionados con el uso del artículo en el corpus de producciones de aprendices estonios de español y especifica claramente cómo computa cada tipo de error:

Los errores de omisión, selección, morfosintácticos y con otros determinantes se calculan con base en el total de grupos no escuetos del corpus (en su versión corregida). Por su parte, los errores de adición se obtienen teniendo en cuenta solo los grupos escuetos (en su versión corregida). (2025: 104)

Además, Rapún (2025) aplica a los resultados análisis estadísticos, en concreto el modelo de regresión logística de efectos mixtos, que permite examinar la evolución de los errores relacionados con el empleo del artículo a lo largo del tiempo.

En nuestro corpus de análisis, Fernández Jódar (2006), Zimny (2006) y Rubinjoni (2017) ofrecen el número de casos correctos, junto con los erróneos, por lo que se puede calcular la proporción de aciertos y de errores (*potential occasion analysis*).

Como en los demás estudios no se facilita ese dato, la comparación entre los resultados no es factible. Solamente Zimny (2006) aplica pruebas estadísticas que determinan la posible correlación entre, por un lado, los aciertos en el uso del artículo y, por otro, la edad de los informantes, el número de horas de estudio de español, el nivel de dominio y el tipo de actividad.

Alguna prueba de fiabilidad —como el coeficiente Kappa (Cohen 1960), calculado por ejemplo por Valverde y Ohtani (2014)— del juicio de gramaticalidad y de la clasificación de los errores —llevados a cabo preferiblemente por más de un investigador y con formación sobre las variedades del español— también es imprescindible y, sin embargo, en los estudios de esta revisión no se menciona ninguna —Rapún (2025: 78) también señala la falta de información sobre la intervención de varios anotadores en el proceso de evaluación en los estudios de análisis de errores que conforman el estado de la cuestión de su tesis doctoral—.

En este sentido, el trabajo de Rapún (2025) es de nuevo un buen modelo, puesto que a los tres especialistas que etiquetan el corpus se les proporciona previamente un esquema de anotación y una taxonomía etiológica de los errores; después se contrastan los resultados del etiquetado de una muestra del corpus y se discuten las divergencias para llegar a un consenso antes de codificar el corpus completo; y, finalmente, también se realiza un análisis estadístico para corroborar que no hay variación significativa entre la codificación de los diferentes observadores. Yonekawa (2025), en un corpus de hablantes de japonés, también presenta un análisis en el que la codificación de los errores la han llevado a cabo dos especialistas, al igual que Zhang (2016) en un corpus de hablantes de chino, y confirma que estadísticamente no hay diferencias significativas entre sus juicios.

El nivel de dominio de los informantes, asimismo, es necesario que se determine objetivamente con alguna prueba de nivel o certificado contrastado. Finalmente, nos parece indispensable contar con un corpus de análisis representativo y que esté informatizado y disponible para poder replicar el análisis o realizar otros. En los trabajos de este análisis, la mayoría de los corpus se incluye mecanografiado e impreso a modo de anexo —véanse Fernández Jódar (2006), Maldonado (2013), Nadal (2012), Rodríguez García (2020), Rubinjoni (2017), Santos Gargallo (1992) y Zimny (2016)—, de ahí que cualquier investigador pueda revisar los textos y analizarlos, gracias a herramientas como #LancsBox X (Brezina y Platt 2025). Al contar con los textos se ha determinado la extensión del corpus¹⁴, cuando el investigador no la indica. De esta manera, se añade precisión a las conclusiones. Los corpus más extensos son los de Rodríguez García (2020) —entre 38 000 y 46 000 palabras—; en segundo lugar, los de Fernández Jódar (2006), Llamas (2020), Maldonado (2013), Montoro (2013), Rubinjoni (2017), Santos

¹⁴ Véanse las notas n.º 4, 5, 6, 8, 9 y 10.

Gargallo (1992), Zimny (2016), que cuentan con entre 14 000 y 28 000 palabras; y, en tercer lugar, los de Castanedo (2019), León, Mestre *et al.* (2017), Lukashenko (2012) y Nadal (2012), con entre 2000 y 6000 palabras.

No obstante, los corpus de textos no siempre están disponibles —Kočman (2011), Llamas (2020), Rascón (2013), León, Mestre *et al.* (2017) y Santiago Alonso (2009) no los facilitan— o solo se ofrece una muestra —como en los estudios de Montoro (2013) y de Castanedo (2019) en los que, además, las producciones están escritas a mano—, lo cual impide la replicabilidad del análisis y reutilizar el corpus para examinar otras cuestiones.

En última instancia, esta investigación pretende contribuir a la mejora de la enseñanza del empleo del artículo a estudiantes de EL2 cuya L1 carece de esta categoría gramatical y facilitar así su aprendizaje.

Bibliografía

- BARALO OTTONELLO, Marta (2004). “La interlengua del hablante no nativo”. En Jesús SÁNCHEZ LOBATO e Isabel SANTOS GARGALLO (Eds.), *Vademécum para la formación de profesores: enseñar español como L2/LE*. Madrid: SGEL, 369-389.
- BELDA TORRIJOS, Mónica (2004). “Análisis de errores ruso-español en estudiantes rusos”. En Milka VILLAYANDRE (Coord.), *Actas del V Congreso de Lingüística General. León 5-8 de marzo de 2002*. Vol. 1. León: Arco/Libros, 357-374.
- BICKERTON, Derek (1995). *Language and human behavior*. Seattle: UW Press.
- BREZINA, Vaclav; PLATT, William (2025). #LancsBox X [software]. Lancaster University. <http://lancsbox.lancs.ac.uk>. [.../.../...]
- CAMPILLOS LLANOS, Leonardo (2012). *La expresión oral en español lengua extranjera: interlengua y análisis de errores basado en corpus*. Tesis doctoral. Universidad Autónoma de Madrid.
- CASTANEDO VEGA, Claudia (2019). *Análisis de errores durante la adquisición del español en rusohablantes de nivel B1 en el aula de ELE: el uso del artículo*. TFM. Universidad de Cantabria.
- CAVADA FERNÁNDEZ-CORONADO, Patricia de la; SESMILO PINA, Álvaro (2007). “Errores específicos en estudiantes eslavos (polacos y croatas) en las destrezas orales”. En Enrique BALMASEDA (Ed.), *Actas del XVII Congreso Internacional de ASELE: las destrezas orales en la enseñanza del español L2-LE*. Logroño: ASELE, 415-430.
- COHEN, Jacob (1960). “A coefficient of agreement for nominal scales”. *Educational and Psychological Measurement*, 20, 37-46.
- CONSEJO DE EUROPA (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación (MCER)*. Madrid: Ministerio de Educación, Cultura y Deporte, y Anaya.

- CORDER, S. Pit (1967). "The significance of learner's errors". *IRAL*, 5(4), 161-170.
- DENST, Edyta (2006). *Uso del artículo en español para hablantes eslavos*. TFM. Universidad Complutense de Madrid.
- ELLER, Helen (2022). *Adquisición/aprendizaje del artículo en español por hablantes de estonio*. Tesis doctoral. Universidad Complutense de Madrid.
- FERNÁNDEZ JÓDAR, Raúl (2006). *Análisis de errores léxicos, morfosintácticos y gráficos en la lengua escrita de aprendices polacos de español*. Tesis doctoral. Universidad Autónoma de Madrid.
- FERNÁNDEZ LÓPEZ, Sonsoles (1991). *Análisis de errores e interlengua en el aprendizaje del español como lengua extranjera*. Tesis doctoral. Universidad Complutense de Madrid.
- FERREIRA CABRERA, Anita; ELEJALDE GÓMEZ, Jéssica (2017). "Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE". *Revista Brasileira de Linguística Aplicada*, 17(3), 509-537.
- HAWKINS, John A.; BUTTERY, Paula (2010). "Criterial features in learner corpora: Theory and illustrations". *English Profile Journal*, 1, 1-23.
- INSTITUTO CERVANTES (2006). *Plan curricular del Instituto Cervantes: niveles de referencia para el español (PCIC)*. Madrid: Biblioteca Nueva.
- KO MAN, Ana (2011). "Errores gramaticales comunes en alumnos croatas de español como segunda lengua en la producción escrita". *marcoELE*, 13, 1-12.
- LEÓN GONZÁLEZ, Blanca; MESTRE MELIÁ, Martí *et al.* (2017). "Analysis of errors by beginner students from Croatia: The case of the definite article in Catalan, Galician and Spanish". *Croatian Journal of Education*, 19(1), 145-155.
- LLAMAS SEVILLA, Jesús M. (2020). *El español como lengua extranjera en nativos de lenguas bálticas orientales: análisis de producciones y estudio comparado de las interlenguas españolas de lituanos y letones*. Tesis doctoral. Universidad Camilo José Cela.
- LUKASHENKO, Nataliia (2012). *Estudio de algunos efectos protolingüísticos en los aprendices ucranianos de español como L2*. TFM. Universidad de Valladolid.
- MALDONADO MARTÍNEZ DE LA CASA, Roberto (2013). *Análisis de errores en la adquisición/aprendizaje del español en estudiantes letones*. TFM. Universidad de Alcalá.
- MARTÍN GONZÁLEZ, Rocío (1999). *El análisis de errores en español como lengua extranjera: el caso del checo*. Tesis doctoral. Universidad Autónoma de Madrid.
- MONTORO CANO, Estela del Rocío (2013). *La enseñanza del español como lengua extranjera en Eslovaquia (Bratislava): historia y actualidad*. Tesis doctoral. Universidad de Granada.
- NADAL LÓPEZ, María José (2012). *Análisis de errores de la interlengua de español para serbiobablantes: dificultades del uso del artículo*. TFM. Universidad de Jaén.
- Biblioteca redELE (2014), 15.

- NIKITINA, Iuliia (2019). *Análisis de errores en el aprendizaje de los artículos en español por parte de aprendientes de habla estonia y rusa*. TFM. Universidad de Tartu.
- RAPÚN MOMBIELA, Virginia (2025). *Evolución en el uso del artículo por parte de estudiantes de grado estonias de lengua española. Un estudio longitudinal de corpus de aprendientes con énfasis en la influencia interlingüística*. Tesis doctoral. Universidad de Tartu.
- RASCÓN CABALLERO, Alfonso (2012). “Análisis de errores e interlengua. Aplicación a las frases con el verbo gustar en estudiantes lituanos de español”. *Verbum*, 3, 91-100.
- RASCÓN CABALLERO, Alfonso (2013). “Avance de análisis de errores de la interlengua en estudiantes lituanos de español”. *Verbum*, 4, 95-105.
- RODRÍGUEZ GARCÍA, Cristina (2020). *Análisis de errores en la interlengua de aprendices de ELE universitarios checos y eslovacos*. Tesis doctoral. Universidad Masaryk.
- RODRÍGUEZ GARCÍA, Cristina (2021). “Errores en el uso del artículo con aprendices de español checos y eslovacos”. *Études romanes de Brno*, 42(1), 381-399.
- RODRÍGUEZ GARCÍA, Cristina (2025). “Errores ortográficos en la interlengua de aprendientes de ELE con L1 checo y eslovaco”. *Revista Internacional de Lenguas Extranjeras*, 24, 231-254.
- RUBINJONI, Vlatka (2017). *Análisis de errores en el uso del artículo determinado en producciones escritas de aprendientes serbios de español como lengua extranjera*. Tesis doctoral. Universidad de Alcalá.
- SANTIAGO ALONSO, Gemma María (2009). “La adquisición del artículo determinado por aprendices eslovenos: análisis cuantitativo y cualitativo de los errores”. *Journal for Foreign Languages*, 1(1-2), 35-51.
- SANTOS GARGALLO, Isabel (1992). *La enseñanza de segundas lenguas, análisis de errores en la expresión escrita de estudiantes de español cuya lengua nativa es el serbo-croata*. Tesis doctoral. Universidad Complutense de Madrid.
- SANTOS GARGALLO, Isabel (2004). “El análisis de errores en la interlengua del hablante no nativo”. En Jesús SÁNCHEZ LOBATO e Isabel SANTOS GARGALLO (Eds.), *Vademécum para la formación de profesores: enseñar español como L2/LE*. Madrid: SGEL, 391-410.
- SANTOS GARGALLO, Isabel; ALEXOPOULOU, Angélica (2021). “Metaanálisis de las tesis doctorales de análisis de errores en la interlengua española a lo largo de tres décadas (1991-2019)”. *marcoELE*, 32, 2-36.
- TARRÉS CHAMORRO, Iñaki (2002). *El uso del artículo por estudiantes polacos de E/LE*. TFM. Universidad de Barcelona.
- TESTA, Martin (2019). “The acquisition of Spanish L3 articles: What can be learned from a simple linear regression analysis?”. *Research in Corpus Linguistics*, 7, 84-112.

- THEWISSEN, Jennifer (2013). "Capturing L2 accuracy developmental patterns: insights from an error-tagged EFL learner corpus". *Modern Language Journal*, 97, 77-101.
- THEWISSEN, Jennifer (2020). "Accuracy". En Nancy TRACY-VENTURA y Magali PAQUOT (Eds.), *The Routledge Handbook of Second Language Acquisition and Corpora*. New York: Routledge, 305-317.
- VALVERDE IBÁÑEZ, María Pilar; OHTANI, Akira (2014). "Annotating article errors in Spanish learner texts: design and evaluation of an annotation scheme". En *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing*, Phuket, Thailand, 234-243.
- VÁZQUEZ, Graciela E. (1991). *Análisis de errores y aprendizaje de español/lengua extranjera*. Frankfurt: Peter Lang.
- YONEKAWA, Kazuko (2025). *Análisis de errores relacionados con el uso del artículo en textos escritos por aprendientes japoneses de español como segunda lengua*. Tesis doctoral. Universidad Nacional de Educación a Distancia.
- ZEČEVIĆ KRNETA, Gorana (2016). *Upotreba određenog člana u španskom kao stranom jeziku na onsovu analize grešaka kod govornika srpskog jezika (Usos del artículo determinado en el ELE en base al análisis de errores de los serbiobablantes)*. Tesis doctoral. Universidad de Kragujevac.
- ZHANG, Yan (2016). *Enseñanza-aprendizaje de ELE y análisis de errores de estudiantes chinos de español*. Tesis doctoral. Universidad de Santiago de Compostela.
- ZIMNY, Ágata (2016). *Análisis de errores en la adquisición del artículo español por alumnos polacos de ELE*. Tesis doctoral. Universidad Complutense de Madrid.