



Claves

La evaluación entre iguales de textos académicos

Un estudio de la actividad metalingüística
del alumnado de grado

Peer evaluation of academic texts

A study of the metalinguistic activity of undergraduate students

SARA GONZÁLEZ BERRIO* | MAITE LÓPEZ FLAMARIQUE**

RESUMEN

El presente trabajo analiza empíricamente la actividad metalingüística de 30 estudiantes de educación superior durante una práctica de evaluación entre iguales de un texto académico. Para ello, se identificaron los errores señalados, se determinó si estaban clasificados correctamente, se comprobó si los comentarios eran correctos, directos o indirectos y se observó si el alumnado había incorporado las sugerencias a la versión final. Los datos muestran que la mayoría de las observaciones se refieren a aspectos superficiales de la escritura, son correctas y se incorporan a las versiones finales. No obstante, el alumnado tiene dificultades para emplear el metalenguaje y utilizar sus conocimientos lingüísticos en el proceso de revisión. El trabajo destaca el potencial de la evaluación entre iguales para desarrollar la actividad metalingüística y la competencia escrita en educación superior, si bien sugiere la necesidad de formar al alumnado en esta práctica e incorporar una gramática pedagógica.

Palabras clave

Evaluación por pares
Escritura académica
Retroalimentación
Metacognición
Educación superior

Recepción: 22 de abril de 2024

Aceptación: 3 de enero de 2025

eLocation: e61776

 <https://doi.org/10.22201/iisue.24486167e.2025.47.61776>

Claves

Abstract

This study empirically analyzes the metalinguistic activity of 30 higher education students during a peer review exercise of an academic text. To this end, the errors identified were noted, it was determined whether they were classified correctly, it was checked whether the comments were correct, direct, or indirect, and it was observed whether the students had incorporated the suggestions into the final version. The data show that most observations refer to superficial aspects of writing, are correct, and are incorporated into the final versions. However, students have difficulty using metalanguage and applying their linguistic knowledge in the revision process. The study highlights the potential of peer assessment to develop metalinguistic activity and written competence in higher education, although it suggests the need to train students in this practice and incorporate pedagogical grammar.

Keywords

Peer assessment
 Academic writing
 Feedback
 Metacognition
 Higher education

* Profesora ayudante doctora de la Universidad Pública de Navarra (España). Doctora en Lingüística Aplicada a la Enseñanza de Lenguas. Líneas de investigación: español como lengua extranjera; español como lengua de acogida; procesos de escritura en L1 y L2; pragmática. Publicaciones recientes: (2023, en coautoría con N. Marcos Miguel), “Tipologías textuales, géneros discursivos y procesos cognitivos en tareas de escritura de Lengua Castellana y Literatura (4º de ESO): ¿cómo integran los libros de texto el currículo oficial?”, *Didáctica. Lengua y Literatura*, vol. 35, pp. 27-41. DOI: <https://doi.org/10.5209/dill.80060>; (2021, en coautoría con S. Martín Leralta y N.A. Rocha), “‘Me es imposible’: elección del género discursivo y estrategias pragmáticas del rechazo en la comunicación académica digital de españoles y brasileños”, *Pragmalingüística*, vol. 29, pp. 90-210. DOI: <https://doi.org/10.25267/Pragmalinguistica.2021.i29.10>. CE: sara.gonzalez@unavarra.es

 ORCID: <https://orcid.org/0000-0003-3006-6126>

** Profesora permanente titular de la Universidad Pública de Navarra (España). Doctora en Innovación e Intervención Educativa. Líneas de investigación: procesos en la escritura colaborativa; alfabetización digital; patrimonio cultural inmaterial en la enseñanza de la lengua. Publicaciones recientes: (2025, en coautoría con I. García-del-Real, M. Aznárez-Mauleón e I. Villarreal), “A Comparative Study of Developing L1 and L2 Writers’ Metalinguistic Activity in Collaborative Writing Tasks among Young Basque Immersion Learners”, *Language Awareness*, vol. 34, núm. 2, pp. 450-475. DOI: <https://doi.org/10.1080/09658416.2024.2419033>; (2022, en coautoría con M. Aznárez-Mauleón e I. García-del-Real), “¿La planificación escrita determina la calidad textual en primaria? Análisis de borradores y producciones finales en una tarea de escritura colaborativa”, *Revista Complutense de Educación*, vol. 33, núm. 1, pp. 1-11. DOI: <https://doi.org/10.5209/RCED.70956>. CE: maite.lopez@unavarra.es

 ORCID: <https://orcid.org/0000-0002-1980-6368>

INTRODUCCIÓN

En los primeros cursos de grado de numerosas universidades se ofrecen talleres de alfabetización académica (Carlino *et al.*, 2013), sin embargo, en estos cursos no siempre se utiliza el potencial formativo y participativo de la evaluación. Lobato Fraile (2006) considera que la evaluación debería ser una parte central del proceso educativo y, al mismo tiempo, una forma de facilitar la regulación y adquisición de contenidos. Desde hace décadas se sabe que la evaluación formativa fomenta el aprendizaje (Scriven, 1967) y que, si el alumnado tiene un papel activo, ésta contribuye a mejorar su formación (Nicol y McCallum, 2022; Panadero *et al.*, 2016). Del mismo modo, se ha comprobado, en el ámbito universitario, que la evaluación participativa y la retroalimentación, vinculadas a las tareas, fomentan el aprendizaje (Rodríguez *et al.*, 2011) y tienen beneficios tanto en el desarrollo de la competencia escrita (Paulus, 1999; Wu y Schunn, 2021) como metacognitiva del alumnado (Rijlaarsam y Couzijn, 1999).

A pesar de las numerosas evidencias que existen sobre los beneficios de la evaluación entre pares, tanto para quien la recibe como para la persona que la realiza (Brown *et al.*, 1994, cit. por Baker, 2016; Dominick *et al.*, 1997; Liu and Carless, 2006; Topping, 1998), su utilización es escasa en el ámbito universitario. De hecho, si bien se emplea una mayor variedad y cantidad de instrumentos de evaluación que hace algunas décadas, todavía predominan los métodos tradicionales, en los que el mayor porcentaje de la nota de la asignatura recae sobre el examen final y la evaluación participativa es casi inexistente (Gómez *et al.*, 2013; Panadero *et al.*, 2019). Es necesario, por tanto, avanzar hacia una evaluación auténtica, participativa, que ofrezca retroalimentación durante el proceso (Ibarra-Sáiz *et al.*, 2020).

En lo que respecta a la evaluación entre iguales, todavía son escasos los trabajos que analizan cómo se desarrolla el proceso y, en concreto, la actividad metalingüística que se produce durante el mismo. En este contexto, el presente trabajo pretende contribuir a cubrir ese vacío al mostrar datos sobre la actividad metalingüística que desarrolla el alumnado universitario durante la evaluación entre iguales de un trabajo académico; esto con el fin de aportar evidencias sobre sus conocimientos del sistema lingüístico y la relación que establecen entre ese saber y el proceso de escritura. Así mismo, pretendemos contribuir al debate abierto sobre la relación existente entre la reflexión gramatical y la mejora del uso lingüístico (Rodríguez Gonzalo, 2012).

Los objetivos específicos de esta investigación son: 1) identificar los errores que los estudiantes corrigen y determinar si identifican de forma correcta el tipo de error en función de distintas categorías preestablecidas (puntuación, aspectos formales, adecuación, cohesión, coherencia, intertextualidad y ortografía); 2) comprobar si los comentarios realizados son correctos y si se

expresan de forma directa o indirecta; y 3) observar si el alumnado incorpora las sugerencias de corrección de sus pares a la versión final del texto.

MARCO TEÓRICO

La escritura en contextos académicos

La incorporación a una institución de educación superior implica, para los futuros estudiantes, formas de actuación y discursos específicos diferentes de los que utilizaron en niveles educativos anteriores (Camps y Castelló, 2013); de ahí la necesidad de iniciar al alumnado de este nivel en las prácticas discursivas de la comunidad académica y, en concreto, en la producción escrita de géneros discursivos propios de las culturas disciplinares (Schneuwly, 2007; Hyland, 2004). En línea con la perspectiva de Bajtín (1992), entendemos los géneros discursivos como enunciados relativamente estables, que varían en función de situaciones típicas de comunicación social y cuyo contenido temático, estilo y composición dependen de cada ámbito. Por lo tanto, generan ciertas expectativas que deben ser cumplidas para participar con éxito en una determinada comunidad discursiva (Reyes, 1999).

La alfabetización académica se puede abordar desde diferentes enfoques no excluyentes: cognitivo, sociocognitivo, sociocultural y socialmente compartido (Castelló *et al.*, 2010a). Desde la perspectiva cognitiva, la redacción se concibe como un proceso recursivo (Flower y Hayes, 1981; Scardamalia y Bereiter, 1992), lo que permite desarrollar estrategias de planificación, textualización y revisión; el enfoque sociocognitivo pone énfasis en la parte motivacional y conductual de la escritura, en interacción con el ambiente físico y social (Zimmerman, 2000); desde la perspectiva sociocultural, la escritura se considera una práctica discursiva, dialógica y situada (Castelló, 2007), en la que la interacción con otros miembros de la comunidad es una fuente de regulación y aprendizaje. La alfabetización académica, así entendida, incluye expectativas e interpretaciones sobre la escritura compartidas por la comunidad académica (Lea y Street, 1998). Una visión de la escritura como práctica socialmente compartida destaca la actividad discursiva como práctica comunitaria (Hadwin *et al.*, 2010) y, tras la aparición de la inteligencia artificial generativa, algunos autores hablan de *ciber-social-learning* (Kalantzis y Cope, 2024), que incluye la interacción con las máquinas en el proceso de escritura.

En concreto, el presente trabajo aborda la alfabetización académica desde las perspectivas sociocultural y cognitiva, y se centra en la monografía como género discursivo perteneciente al ámbito académico. De acuerdo con Carlino (2005), se trata de un análisis bibliográfico escrito para ser evaluado y consta de cuatro partes: introducción, desarrollo, conclusiones y referencias.

La introducción consiste en presentar el tema, el desarrollo debe defender una idea a partir del análisis de un conjunto de textos, las conclusiones sintetizan el argumento principal y las referencias muestran las fuentes consultadas y citadas.

La evaluación entre iguales

De los diversos tipos de evaluación participativa (autoevaluación, evaluación entre iguales y coevaluación), en nuestro trabajo nos centramos en la evaluación entre iguales (*peer assessment* o *peer review*) que, al ser también formativa, cumple una doble función: por un lado, promueve procesos de autorregulación del aprendizaje y, por otro, desarrolla la responsabilidad del alumnado con su propio proceso de aprendizaje y el de sus semejantes. Si bien cada tipo de evaluación participativa tiene sus propias características, ventajas y limitaciones, todas ellas han demostrado ser eficaces para fomentar tanto el aprendizaje como las estrategias de autorregulación (Nicol y McCallum, 2022; Panadero *et al.* 2016; Liu y Carless, 2006, entre otros).

En la literatura a menudo conviven como sinónimos los términos “evaluación entre iguales” y “coevaluación”; no obstante, preferimos seguir la pauta de Gómez y Quesada (2017) y distinguir ambas modalidades. De ahí que entendamos la coevaluación como aquella que “es realizada y negociada de forma conjunta entre docentes y estudiantes, en la que la responsabilidad es compartida, en la que es imprescindible el diálogo y en la que se prima llegar a un consenso sobre las valoraciones realizadas” (Gómez y Quesada, 2017: 11). Por su parte, la evaluación entre iguales, según Rodríguez *et al.* (2013: 202), es el “proceso mediante el cual los estudiantes realizan un análisis y valoración sobre las actuaciones y/o producciones desarrolladas por algún estudiante o grupo de estudiantes de su mismo estatus o nivel”.

La evaluación entre iguales ha resultado eficaz para la mejora de la escritura en la universidad, tanto en L1 como en L2 (Bueno-Alastuey *et al.*, 2022; Rodríguez *et al.*, 2011; Lundstrom y Baker, 2009; Topping, 1998, entre otros). Un metaanálisis realizado por Huisman *et al.* (2019) concluye que ésta logra mejores resultados en la calidad de los textos que la autoevaluación o ausencia de *feedback*, y similares efectos a la evaluación del profesorado. En esta misma línea, existen numerosas evidencias de que, al integrar, aunque sea de forma parcial, las observaciones y los comentarios realizados durante el proceso de evaluación del texto, se obtiene una mejora tanto del escrito final (Cho y Cho, 2011; Dressler *et al.*, 2019, entre otros), como de la competencia escrita (Paulus, 1999) y beneficia, especialmente, a quienes presentan más problemas de redacción (Wu y Schunn, 2021).

En cuanto al análisis de esta práctica evaluativa, Alshuraidah y Storch (2019) observaron que 79 por ciento del *feedback* proporcionado por estudian-

tes adultos se refería a aspectos superficiales de la lengua, 11 por ciento a la estructura y 18 por ciento al contenido del trabajo. De los aspectos lingüísticos, más de la mitad (65 por ciento) eran cuestiones de puntuación y ortografía. Un estudio realizado por Castelló *et al.* (2010) analiza un proceso de evaluación entre pares con estudiantes de posgrado y muestra que más de la mitad de los problemas identificados por ellas y ellos están relacionados con la cohesión de frases y párrafos (uso de conectores, sintaxis y falta de progresión temática). En el mismo sentido, Yu (2021) halló que la mayor parte del *feedback* realizado entre estudiantes de posgrado se centra en aspectos lingüísticos, de contenido y organización de la tesis.

A la hora de incorporar las propuestas de corrección recibidas, Dressler *et al.* (2019) observaron que el alumnado integró 85 por ciento del total de ítems de corrección recibidos. Los comentarios referidos a aspectos superficiales de la escritura se incorporaron en mayor medida (89.2 por ciento) que aquéllos relativos al contenido del texto (82.9 por ciento). Estos autores señalan que esta diferencia puede ser debida a que el cambio de contenido exige mayor trabajo al estudiante que recibe la corrección. Asimismo, la falta de una definición clara de los problemas encontrados y los comentarios poco precisos pueden ser un impedimento para la mejora del texto (Castelló *et al.*, 2010b).

Por último, es imprescindible hacer hincapié en la necesidad de formar al alumnado para que realice una evaluación efectiva (Baker, 2016). De hecho, se ha observado que recibir formación sobre cómo proporcionar *feedback* incrementa la capacidad de realizar una evaluación más efectiva, que aborde no sólo aspectos superficiales de escritura, sino también problemas de contenido (Zhao, 2014, cit. por Alshuraidah y Storch, 2019).

La actividad metalingüística durante la revisión entre iguales

La actividad metalingüística puede ser definida como aquella que es observable, verbal o procedimental y que toma el lenguaje como objeto de reflexión y análisis (Camps y Milian, 2000). Según advierte Castelló (2000), el análisis de este tipo de actividad nos permite acceder al conocimiento previo del alumnado sobre los diferentes niveles lingüísticos, desde el código gráfico hasta la gramática oracional y textual. Dicho conocimiento sobre la lengua puede ser declarativo, es decir, dominar tanto las reglas y los principios lingüísticos como la habilidad para enunciarlos; y procedimental, esto es, ser capaz de poner dicho conocimiento en acción (Gombert, 1992). Dado que este tipo de actividad puede ser más o menos consciente, se puede hablar de un continuum que va, desde el uso de la lengua que se produce de manera automática en un extremo, hasta la reflexión metalingüística realizada de manera explícita y, en ocasiones, con terminología específica, en el otro (Tolchinsky, 2000). En este

sentido, algunas autoras (Tolchinsky, 2000; Hugo y Meneses, 2021) diferencian entre actividad metalingüística, aquella referida al hecho lingüístico en general con cualquier nivel de conciencia, y reflexión metalingüística, la que presta atención explícita a determinados rasgos lingüísticos o discursivos y que exige mayor conciencia sobre el sistema de la lengua.

Habida cuenta del grado de conciencia y de conocimientos lingüísticos que exige la práctica de evaluación entre iguales, nuestra propuesta estaría más cerca de la reflexión metalingüística; no obstante, se optó por emplear el término “actividad metalingüística” en su acepción más amplia, entendida como un continuum que incluye operaciones cognitivas más automáticas —como analizar el sentido global de un texto— y otras más conscientes como, por ejemplo, clasificar los errores identificados.

Si bien se ha observado que la actividad metalingüística se da a lo largo de todo el proceso de escritura (Flower y Hayes, 1981; Aznárez-Mauleón *et al.*, 2020), la revisión de lo escrito favorece la reflexión sobre los conocimientos lingüísticos y la toma de conciencia sobre la lengua (Rijlaarsam y Couzijn, 1999; Liu y Carless, 2006; Yarrow y Topping, 2001). El alumnado debe identificar los errores o problemas que tiene el texto y construir un razonamiento explicativo a partir de una reflexión metalingüística (Liu y Carless, 2006, Topping, 1998). Esto requiere realizar operaciones cognitivas de alto nivel (Boud, 1990), como la discriminación entre correcto e incorrecto, la evaluación global del texto o el reconocimiento de las características de los géneros textuales (Wu y Schunn, 2021). Yu (2021) encontró que 85 por ciento de los comentarios realizados por los estudiantes, tanto aquéllos centrados en aspectos lingüísticos como los referidos a las características del género, eran correctos. No obstante, existen posibles limitaciones del *feedback* entre iguales, pues cabe la posibilidad de que la persona que realiza la evaluación del escrito no posea la capacidad evaluativa suficiente (Pitt *et al.*, 2020; Hill y West, 2020, cit. por Cano *et al.*, 2022) o el conocimiento lingüístico necesario (Yu, 2021). De hecho, Schere (2020) encontró que buena parte de las y los estudiantes de los primeros cursos universitarios tienen dificultades relacionadas con la sintaxis de las oraciones, lo cual condiciona el resultado de sus producciones escritas.

Otra de las ventajas que presenta la evaluación entre iguales es que, al revisar el texto de otra persona, se tiene la oportunidad de reparar en otros modelos de escritura y se puede tomar conciencia de algunos errores de la propia (Boud, 1995; Rijlaarsam y Couzijn, 1999). Lundstrom y Baker (2009) también concluyeron que quienes evalúan otros textos, autorregulan y mejoran su propia escritura mediante la observación y el aprendizaje mimético. Es decir, estas acciones fomentan la conciencia lingüística y la actividad metalingüística (Liu y Carless, 2006; Sanmarti, 2007).

En definitiva, la alfabetización académica es necesaria para acercar al alumnado a las distintas prácticas discursivas del ámbito universitario y una

forma de hacerlo es a través de los distintos géneros discursivos. Se ha visto que la evaluación entre iguales promueve la mejora de la competencia escrita y contribuye a la reflexión sobre el uso de la lengua, si bien son pocas las investigaciones que han realizado un análisis sistemático de esta actividad. En este estudio se analiza la actividad metalingüística subyacente a la evaluación entre pares de una monografía con el fin identificar el conocimiento del alumnado sobre diferentes aspectos del sistema lingüístico. A continuación, se describe pormenorizadamente tanto la intervención realizada como el método empleado para la recogida de datos y el posterior análisis.

MÉTODO

El presente estudio se enmarca en el paradigma cualitativo de investigación, enfoque adecuado para comprender los fenómenos didácticos (Bogdan y Biklen, 1997). En concreto, responde a un estudio de caso (Yin, 2009; Simons, 2011; Yazan, 2015), pues analiza empíricamente un fenómeno educativo dentro de su contexto real. El diseño metodológico escogido se justifica porque nuestras preguntas de investigación pretenden estudiar el “cómo” y, para ello, la respuesta exige una descripción exhaustiva (Yin, 2009). Según explica Stake (1999), en el ámbito educativo es habitual que los estudios de caso se realicen con un grupo de individuos, tal y como ocurre en el presente trabajo. De esa forma, se lleva a cabo un análisis individual y colectivo con el fin de identificar consistencias e inconsistencias (Yin, 2009).

La toma de datos de esta investigación se realizó en el marco de la asignatura obligatoria Habilidades comunicativas y TIC, impartida en el primer año de los grados de Educación de una universidad pública española. En concreto, en este estudio participaron 30 estudiantes, 24 mujeres y 6 hombres (edad media de 18 años 2 meses) de primer curso de los grados de Educación infantil y Educación primaria.

Entre los objetivos de esta asignatura está el aprendizaje de la elaboración de una monografía, de forma individual, sobre un tema de libre elección, pero relacionado con la educación. En la asignatura, esta revisión bibliográfica recibe el nombre de “trabajo académico”, por lo que vamos a utilizar ambos términos como sinónimos. La intervención didáctica, de seis semanas de duración con dos sesiones semanales, se organizó en función de las diferentes operaciones del proceso de escritura (Cassany, 1999) y de los conocimientos lingüísticos y discursivos que caracterizan la monografía. Tanto la adopción de un género propio de la universidad como la evaluación entre pares contribuyen a que el aprendizaje sea significativo y socialmente situado (Navarro, 2018).

Durante la planificación de la monografía se realizaron actividades para la elección del tema, la búsqueda de artículos científicos y se dieron pautas tanto de

lectura como de vaciado de información. La actividad metalingüística se centró en la identificación de las características lingüísticas y discursivas propias de la escritura académica, y en la selección de ideas. Como resultado, el alumnado presentó un proyecto inicial de su trabajo que recibió *feedback* de la docente.

A lo largo del proceso de textualización se realizaron actividades para trabajar la coherencia y cohesión textual, los signos de puntuación, la adecuación y el uso de un lenguaje inclusivo. Para evaluar formativamente esta parte, la escritura de la introducción se realizó en el aula y recibió retroalimentación de la docente.

Antes de la entrega final de la monografía, las y los estudiantes participaron en un proceso de evaluación entre iguales que tiene como objetivo promover la reflexión consciente sobre el uso de la lengua en la escritura y la mejora del texto escrito. El emparejamiento del alumnado para la realización de la evaluación se realizó a través de la plataforma virtual de la asignatura que vincula a las y los participantes de forma aleatoria. Este proceso de evaluación consta de las siguientes etapas:

1. Escritura del borrador.
2. Intervención didáctica sobre cómo realizar la evaluación de las monografías.
3. Asignación aleatoria de monografías para su evaluación entre iguales.
4. Evaluación del trabajo de una compañera o un compañero mediante comentarios realizados con el procesador de textos Microsoft Word.
5. Devolución del trabajo evaluado.
6. Redacción de la versión final tomando en cuenta el *feedback* recibido.
7. Evaluación de las monografías por parte de la docente.

La intervención didáctica sobre la evaluación entre pares consiste en realizar un ejercicio similar en el aula, primero de forma individual y más tarde con la docente. Para ello, se proporcionó al alumnado una monografía de cursos anteriores, un documento con las instrucciones y una clasificación de errores (Tabla 1), adaptada de Cassany (1993). Se pidió al alumnado que seleccionara los fragmentos en los que considerara que había algún error y que incluyera un comentario, utilizando la herramienta del procesador de textos Word; y, además, que indicara el tipo de error y la explicación. Se les recordó que podían consultar el material de clase y diferentes fuentes de consulta en línea, como el *Diccionario de la lengua española*, la Fundación del Español Urgente (Fundéu) y el *Diccionario Panhispánico de Dudas*, previamente utilizados en las sesiones presenciales. En la corrección conjunta se comentaron las dudas que surgieron. Con estas mismas pautas, el alumnado debería evaluar otro trabajo que recibiría a través de la plataforma *online* de la asignatura.

Tabla 1. Aspectos a revisar en la evaluación entre pares

Aspecto	Descripción
Ortografía	Errores ortográficos, uso incorrecto de tildes o mayúsculas
Puntuación	Errores en el uso de los signos de puntuación o de los signos auxiliares
Adecuación	El registro no es formal, no se utiliza un lenguaje inclusivo, se utilizan palabras comodín, se repiten innecesariamente algunos términos...
Cohesión	Errores de concordancia, uso incorrecto de conectores o preposiciones, ordenación sintáctica inadecuada, idea que no está bien expresada...
Coherencia	Falta de coherencia entre las diferentes partes del texto, contenido inadecuado en el apartado, falta de progresión temática, desorganización de información, uso incorrecto de marcadores textuales...
Intertextualidad	No se cita la fuente, no se siguen las normas APA o no se incluyen artículos científicos (al menos tres)
Aspectos formales	Se respetan los requisitos formales del trabajo

Fuente: elaboración propia, adaptada de Cassany, 1993.

Es importante destacar que la práctica de evaluación entre iguales analizada se realiza fuera del aula y el alumnado tiene tres días para entregarla. Una vez recibida su monografía con comentarios de corrección, disponen de dos días para incorporar las correcciones que consideren oportunas y hacer la entrega de la versión final del trabajo. En este sentido, se les advierte que deben considerar la conveniencia o no de realizar los cambios sugeridos, pues cabe la posibilidad de que no siempre sean correctos.

Para el presente estudio se recogieron los datos obtenidos a través de los borradores (n=30) y las versiones finales (n=30). En el primer caso, se analizaron los comentarios al margen realizados durante la práctica de evaluación entre iguales; en el segundo, se cotejó la versión final con el borrador para comprobar si se incorporaron las propuestas de corrección. Junto a la entrega del trabajo, se pidió al alumnado el consentimiento informado para participar en esta investigación, de acuerdo con las recomendaciones del Comité de Ética de la universidad donde se realizó la toma de datos.

El análisis de cada uno de los comentarios se realizó mediante una codificación por capas, como figura en la Tabla 2. En una primera capa se codificó el nivel lingüístico o discursivo al que correspondía la corrección; para ello se tomó como referencia el trabajo de Aznárez-Mauleón *et al.* (2020) y se incluyeron dos nuevas categorías de acuerdo con los requisitos de la tarea: intertextualidad (utilización adecuada de las normas APA al citar y referenciar) y aspectos formales (tipo de letra, tamaño...). Cabe puntualizar que, en la categoría de intertextualidad, se incluyeron los criterios formales relacionados con las normas de citación, tales como el uso de sangría y la utilización de comillas. En una segunda capa se etiquetó si la clasificación del tipo de error identificado era correcta, incorrecta, o si, simplemente, no se indicaba. En un

tercer momento se codificó si el *feedback* ofrecido era correcto, incorrecto o estilístico. Y, por último, se indicó si el comentario era directo o indirecto. Posteriormente, se comparó el borrador del trabajo con su versión final y se identificó si los cambios sugeridos habían sido incorporados a la versión final.

Para la codificación de los datos se utilizó el programa de análisis cualitativo Atlas.ti (versión 8), lo que permitió cuantificar la frecuencia de aparición de los ítems estudiados para los posteriores análisis. Con el fin de asegurar la fiabilidad de la codificación, una parte del corpus (5 borradores, 16 por ciento aproximadamente) fue codificada por las dos autoras del estudio. Se comparó el etiquetado y se discutieron las diferencias hasta llegar a un acuerdo. El resto del corpus se dividió entre las dos investigadoras para su codificación y, en caso de duda, se tomaron decisiones conjuntas.

Tabla 2. Categorías de análisis de la actividad metalingüística

Capas de codificación	Categoría	Aspecto	Definición
1ª capa	Nivel lingüístico y discursivo	Ortografía	Ver Tabla 1
		Puntuación	
		Adecuación	
		Cohesión	
		Coherencia	
		Intertextualidad	
		Aspectos formales	
2ª capa	Clasificación del tipo de error	Correcta	Se identifica correctamente el tipo de error
		Incorrecta	Se identifica incorrectamente el tipo de error
		Sin clasificar	No se menciona el tipo de error
3ª capa	Comentario	Correcto	El comentario responde a criterios lingüísticos
		Incorrecto	El comentario no responde a criterios lingüísticos
		Estilo	Tanto la versión del texto original como la observación realizada son válidas porque ambas responden a criterios lingüísticos
4ª capa	Tipo de comentario	Directo	El comentario incluye la solución
		Indirecto	El comentario indica que hay un error y lo justifica, pero no incluye la solución
5ª capa	Modificaciones en la versión final	Incorporado	El comentario recibido durante la evaluación entre iguales se incorporó a la versión final
		No incorporado	El comentario recibido durante la evaluación entre iguales no se incorporó a la versión final

Fuente: elaboración propia.

RESULTADOS

A continuación, se presentan los resultados obtenidos tras analizar tanto los trabajos revisados como las versiones finales. Se presentan, en primer lugar, los errores identificados durante el proceso de evaluación entre iguales y cómo los clasifica el alumnado. En segundo lugar, se describe si se proponen cambios correctos, incorrectos o si se trata de cuestiones estilísticas. Por último, se detalla el tipo de retroalimentación ofrecida (directa o indirecta) y se analiza si las modificaciones propuestas por sus pares son incorporadas a la versión final del trabajo académico.

Errores identificados durante la evaluación entre iguales

La mayor parte de los comentarios realizados durante la evaluación entre iguales están relacionados con la puntuación (24.7 por ciento). Destacan también las observaciones sobre aspectos formales (16.5 por ciento), adecuación (14.9 por ciento), cohesión (14.9 por ciento), intertextualidad (14.3 por ciento) y, en menor medida, ortografía (10.9 por ciento). En contraposición, la categoría que recibió menos apreciaciones fue la coherencia (3.81 por ciento). A continuación, la Tabla 3 muestra los errores identificados en cada categoría.

Tabla 3. Tipo de error identificado

Tipo de error	Núm. (%)	\bar{x} (σ)	Ejemplo
Puntuación	136 (24.7)	4.5 (4)	Sobran los dos puntos Falta el punto final
Aspectos formales	91 (16.5)	3 (4.2)	Tiene que estar en [letra] 14 Los márgenes deben ser de 3 cm
Adecuación	82 (14.9)	2.7 (4)	El registro no es formal. Es una expresión coloquial No se utiliza un lenguaje inclusivo
Cohesión	82 (14.9)	2.7 (3.8)	No está bien expresada la idea Orden sintáctico inadecuado No se entiende la frase
Intertextualidad	79 (14.3)	2.6 (3.4)	Este texto tiene que ir en la bibliografía
Ortografía	60 (10.9)	2 (2.8)	Sin tilde A gusto
Coherencia	21 (3.8)	0.7 (1.1)	Repetición de la idea anterior Debería ser otro párrafo
Total	551 (100)	33.9 (20.5)	

Fuente: elaboración propia.

A propósito de los comentarios sobre puntuación, cabe destacar que, de las 135 anotaciones, sólo en 13.5 por ciento de ellas se da una explicación de por qué

hay que realizar el cambio, es decir, se ofrece una reflexión metalingüística explícita. En cinco de ellas se plantea que “no es necesario poner coma entre el sujeto y el verbo”; otras tantas aluden a que los marcadores o conectores deben ir entre comas y el resto se refieren a otros aspectos, como que el título no lleva punto o la necesidad de poner punto final a las oraciones. El 86.5 por ciento de los comentarios se limita a señalar que falta o sobra algún signo de puntuación; son las comas el signo que centra la atención del alumnado, si bien hay algunas referencias a un uso inadecuado del punto y los dos puntos. Los comentarios sobre aspectos formales (91) se refieren principalmente a que el texto no se ajusta al tamaño o tipo de letra requeridos o no está bien justificado. En ocasiones, también se indica que sobran espacios entre las palabras.

En cuanto a los 82 errores de adecuación encontrados, que engloban también los comentarios referidos a un uso inadecuado del vocabulario, 40 (48.7 por ciento) son advertencias de que no se está utilizando lenguaje inclusivo o no se hace de forma adecuada (uso de @ o x, por ejemplo); el resto se refiere a la repetición de términos o la utilización de coloquialismos. En el primer caso, algunas veces el comentario es muy explícito: “repites ‘de manera’ en la misma frase, busca sinónimos”; en otros, solamente se alude al problema encontrado: “pobreza léxica, repetición”. Si nos fijamos en la cohesión, se hace referencia a aspectos concretos como el uso de conectores, la utilización errónea de preposiciones (“en vez de ‘en’, pondría ‘de’”, “respecto a”) o la falta de concordancia entre sujeto y verbo (“tenga, profesorado es singular”).

Sobre la categoría de intertextualidad (79), la mayoría de los comentarios (51) hacen referencia al incumplimiento de las normas APA en las citas: “falta la página”, “sólo apellido”, “no hace falta poner día y mes”; “la cita deber ir entrecomillada”, “al ser una cita de más de 40 palabras debe ir sangrada”. En menor medida, hay anotaciones sobre aspectos formales que deben cumplir las referencias bibliográficas (12), la importancia de que todas las fuentes citadas aparezcan en la bibliografía (8) y la necesidad de que las referencias aparezcan en orden alfabético (7). De los 60 comentarios sobre ortografía, algo más de la mitad (31) se refiere al uso de la tilde; en menor medida hay sugerencias relacionadas con la utilización de mayúsculas o minúsculas (8) y también advierten sobre palabras que están mal escritas como “agusto”, “a parte” (en referencia al adverbio aparte) o “microracismos” (6), si bien estos casos son más puntuales. Por último, los comentarios referidos a la coherencia del texto (21), en su mayoría, advierten de la repetición de ideas (12) o de aspectos contradictorios (6); también aluden a que el contenido no se ajusta a lo solicitado (6), por ejemplo, “falta la justificación de la elección del tema: motivación personal” o “en la conclusión podrías mencionar las soluciones”. Algunos comentarios sugieren la necesidad de organizar la información de otra manera: “debería ir en otro párrafo” (4) y otros piden más concreción: “¿a qué te refieres?”, “¿qué tipo de educación?” (3). Al fijarnos en los datos que nos

aporta la estadística descriptiva, llama la atención que, a pesar de haber una media de 4.5 comentarios referidos a la puntuación en cada trabajo, también es uno de los aspectos que presenta una mayor desviación estándar; es decir, hay quienes incluyeron numerosas observaciones y quienes apenas repararon en ella. También es alta la desviación estándar en el caso de los aspectos formales (4.2) y la adecuación (4).

El mayor porcentaje de aciertos al clasificar el tipo de error corresponde a la puntuación (72.8 por ciento), seguido de la ortografía (61.7 por ciento) y la intertextualidad (60.8 por ciento), como figura en la Tabla 4. En cambio, los comentarios relacionados con la cohesión (24.4 por ciento) y la adecuación (14.6 por ciento) fueron clasificados incorrectamente en más ocasiones que el resto de los aspectos analizados. Asimismo, hay numerosos errores identificados que no han recibido ningún tipo de clasificación. Entre ellos destacan aquéllos vinculados a aspectos formales (55 por ciento) y a la adecuación (40.2 por ciento).

Tabla 4. Clasificación del tipo de error

Tipo de error identificado	Clasificación correcta		Clasificación incorrecta		Sin clasificación	
	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)
Puntuación	99 (72.8)	3.3 (3.9)	11 (8.1)	0.4 (1.4)	26 (19.1)	0.9 (1.9)
Aspectos formales	36 (39.6)	1.2 (1.9)	5 (5.5)	0.2 (0.5)	50 (55)	1.7 (4)
Adecuación	37 (45.1)	1.2 (2)	12 (14.6)	0.4 (0.7)	33 (40.2)	1.1 (2.9)
Cohesión	48 (58.5)	1.6 (3)	20 (24.4)	0.7 (1.1)	14 (17.1)	0.5 (0.9)
Intertextualidad	48 (60.8)	1.6 (3.4)	3 (3.8)	0.1 (0.3)	28 (35.4)	0.9 (1.7)
Ortografía	37 (61.7)	1.2 (2.4)	3 (5)	0.1 (0.3)	20 (33.3)	0.7 (1.7)
Coherencia	11 (52.4)	0.4 (0.8)	1 (4.8)	0 (0.2)	9 (42.9)	0.3 (0.7)
Total	315 (100)	10.5 (12.8)	55 (100)	1.8 (2.4)	180 (100)	6 (7.4)

Fuente: elaboración propia.

Corrección de los comentarios realizados durante la evaluación entre iguales

La mayoría de las observaciones que realiza el alumnado en los textos elaborados por sus iguales son correctas, especialmente las que se refieren a los aspectos formales (87.9 por ciento) y la ortografía (85 por ciento); sin embargo, algunos comentarios realizados durante este tipo de evaluación son incorrectos. En el caso de la intertextualidad, un 25.3 por ciento no son correctos. Por ejemplo, se anota “sobran las comillas” cuando se trata de una cita textual de menos de 40 palabras o se comenta “cuando son más de dos autores, se pone et al.” junto a una cita múltiple de textos diferentes. A continuación, en la Tabla 5 figuran los distintos tipos de corrección por categorías.

Tabla 5. Tipo de corrección: correcto, incorrecto o de estilo

Tipo de error identificado	Comentario correcto		Comentario incorrecto		Cuestión estilística	
	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)
Puntuación	97 (71.3)	3.2 (2.9)	24 (17.6)	0.8 (1)	15 (11)	0.5 (1)
Aspectos formales	80 (87.9)	2.6 (4)	6 (6.6)	0.2 (0.6)	5 (5.5)	0.2 (0.4)
Adecuación	65 (79.3)	2.2 (3.4)	8 (9.8)	0.3 (0.6)	9 (11)	0.3 (0.7)
Cohesión	53 (64.6)	1.8 (2.9)	10 (12.2)	0.3 (0.6)	19 (23)	0.6 (1.7)
Intertextualidad	55 (69.6)	1.8 (3.2)	20 (25.3)	0.7 (1.2)	4 (5)	0.1 (0.6)
Ortografía	51 (85)	1.7 (2.6)	8 (13.3)	0.3 (0.8)	1 (1.7)	0 (0.2)
Coherencia	14 (66.7)	0.5 (0.9)	5 (23.8)	0.2 (0.5)	2 (9.5)	0.1 (0.3)
Total	413 (100)	13.8 (12.7)	81 (100)	2.7 (2.5)	55 (100)	1.8 (3.4)

Fuente: elaboración propia.

La puntuación, al tener el mayor número de comentarios, presenta el número más alto de anotaciones incorrectas: 24 (17.6 por ciento de los totales). La mayoría se derivan de un conocimiento sintáctico erróneo, como no reconocer el sujeto de la oración. A modo de ejemplo, en uno de estos casos un alumno pide que se quite la coma del siguiente pasaje, alegando que no hay que poner coma entre sujeto y verbo: “Con este texto, se puede concluir...”. En otros comentarios, se solicita poner un punto en mitad de una idea o no se identifica bien el conector que debe ir entre comas.

En el caso de la cohesión, algunas anotaciones incorrectas denotan un uso erróneo de las preposiciones: una alumna pide que se añada una *a* tras la palabra *sexo* en el siguiente fragmento: “Se entiende por *sexo* las diferentes características biológicas...”. Otras simplemente responden a una deficiente comprensión del texto. Encontramos comentarios como: “no se entiende bien”, “la idea está mal expresada”, cuando, en realidad, el texto está escrito correctamente.

Asimismo, se identificaron comentarios que, lejos de responder a normas lingüísticas prescriptivas, incluyen cuestiones de estilo y, por tanto, son facultativos. Esto es más evidente en cuestiones de cohesión (23.2 por ciento), pero también es palpable en la puntuación (11 por ciento) y la adecuación (11 por ciento). Un ejemplo es proponer a un compañero que, en vez de *porque*, ponga *ya que*. En otro caso, se solicita reemplazar *a día de hoy* por *hoy en día*. Otras anotaciones hacen referencia a comas que no son prescriptivas, sino meras cuestiones de estilo.

Al fijarnos en la dispersión de los datos con respecto al promedio, observamos que es mayor en el caso de los comentarios correctos y, en especial, de los aspectos formales. Esto quiere decir que, si bien el alumnado realizó una media de 2.6 comentarios correctos sobre los aspectos formales, son numerosos quienes llegaron a realizar 6.6 o ninguno. En cambio, la tendencia en el caso de los comentarios incorrectos y las cuestiones estilísticas es más constante.

Tipo de retroalimentación ofrecida durante la evaluación entre iguales

La gran mayoría de los comentarios que indican explícitamente la forma correcta (comentarios directos) hacen referencia a la puntuación (35.1 por ciento) y, en menor medida, a la cohesión (16.4 por ciento), los aspectos formales (15.5 por ciento) y la ortografía (14.9 por ciento). Tal y como se puede ver en la Tabla 6, en el caso de las observaciones que recogen una explicación del error (comentarios indirectos), los aspectos que reciben mayor atención son los de adecuación (27 por ciento), intertextualidad (19.9 por ciento), aspectos formales (18 por ciento) y, de nuevo, la cohesión (12.9 por ciento).

Cabe mencionar que algunos comentarios incluyen tanto la explicación del error como la forma correcta; de ahí que las etiquetas no sean excluyentes. Por eso, en algunos casos, el número total de comentarios directos e indirectos es superior al número relativo a cada tipo de error. Por ejemplo, uno de los comentarios dice: “uso incorrecto del conector. Quita el *pero* y pon una coma”. En este caso se da tanto la explicación directa como la indirecta.

En cuanto a los aspectos formales, en la mayor parte de los casos explícitamente se señala cómo se debe realizar la corrección: “tienes que poner tamaño 14”, “falta sangría”. En los casos de corrección indirecta se recuerda el criterio solicitado: “no utilizas la letra que se pide”. Algo similar ocurre con los comentarios sobre intertextualidad: los directos apuntan el fallo concreto (“sobra el nombre del autor”, “falta la página”), mientras que los indirectos aluden a las normas de referencia (“revisa normas APA”).

Tabla 6. Tipo de comentario: directo o indirecto

Tipo de error identificado	Comentarios directos		Comentarios indirectos	
	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)
Puntuación	122 (35.1)	4.1 (4.1)	22 (8.6)	0.7 (1.1)
Aspectos formales	54 (15.5)	1.8 (3.5)	46 (18)	1.5 (2.2)
Adecuación	24 (6.9)	0.8 (1.4)	69 (27)	2.3 (3.7)
Cohesión	57 (16.4)	1.9 (3.2)	33 (12.9)	1.1 (1.6)
Intertextualidad	34 (9.8)	1.1 (2)	51 (19.9)	1.7 (2.4)
Ortografía	52 (14.9)	1.7 (2.8)	17 (6.6)	0.6 (1.2)
Coherencia	5 (1.4)	0.2 (0.4)	18 (7)	0.6 (1.1)
Total	348 (100)		256 (100)	

Fuente: elaboración propia.

Si nos fijamos en las anotaciones referidas a la cohesión, en 57 ocasiones el alumnado escribió directamente en el comentario la forma que considera correcta, sin dar ningún tipo de explicación de por qué es necesario realizar

el cambio. De los comentarios en los que proporcionan una explicación del error detectado, cabe destacar que casi la mitad de ellos (14 de 33) resultan muy inespecíficos: “la idea está mal expresada”; “suena rara la frase, se hace difícil leerla”, “no se entiende la frase”, “reformulación de la frase, no tiene sentido”. Esto denota falta de vocabulario metalingüístico y dificultades para utilizar sus conocimientos lingüísticos en el proceso de revisión.

En las anotaciones sobre puntuación sucede algo similar. El grueso de los comentarios responde a una corrección directa: se limitan a señalar si se debe añadir o eliminar algún signo de puntuación, principalmente, comas. Entre las correcciones indirectas, al dar la explicación del fallo, en ocasiones, la idea no es del todo correcta: “no se pone coma delante del verbo”. Sólo en algunos comentarios se utiliza correctamente el metalenguaje (“los marcadores van entre comas”) y, en escasas ocasiones, se emplea de forma muy precisa: “los conectores consecutivos se escriben entre comas”.

Con relación a la adecuación, los comentarios indirectos apuntan el tipo de fallo: “uso de coloquialismo”, “repetición de conectores”, “no usas lenguaje inclusivo”. Los directos, en cambio, realizan propuestas concretas: “niños y niñas”, “alumnado”. En algunos casos, como hemos señalado al inicio, se incluyen los dos tipos de corrección: “Se repite la palabra, pon *extraer*”.

Para terminar, en lo que respecta a la coherencia textual, la mayoría de las anotaciones son indirectas (“se repite la idea”, “falta la motivación personal”) y, en muy pocos casos (5) se sugiere añadir o suprimir algo del texto original.

Incorporación de las propuestas de corrección a la versión final

Al analizar la incorporación de las propuestas de corrección a la versión final del texto (Tabla 7) se tuvo en cuenta el tipo de error al que hacen referencia, si la propuesta realizada es correcta y si está formulada de forma directa (proporciona la solución) o indirecta (indica el camino para llegar a una solución). *Grosso modo*, la mayoría de los comentarios incorporados corresponde a la puntuación (27 por ciento) y a los aspectos formales (17 por ciento). En contrapartida, son muy escasos los comentarios sobre coherencia (2 por ciento) que se llegan a incorporar a los trabajos escritos. En parte, puede ser porque la mayoría de ellos se formulan de manera indirecta.

Al observar los comentarios no incorporados a la versión final, los que se refieren a cuestiones de intertextualidad (22 por ciento) destacan por encima de los demás. No obstante, tiene sentido porque, como hemos mencionado con anterioridad, esos comentarios son los más incorrectos y, además, la mayoría son indirectos. Entre los comentarios no incorporados les siguen aquéllos relacionados con la adecuación (17 por ciento), los aspectos formales (16 por ciento) y la puntuación (16 por ciento). Al analizar la dispersión de los datos

Tabla 7. Comentarios incorporados y no incorporados por categorías

Aspecto comentado	Comentario incorporado		Comentario no incorporado	
	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)
Puntuación	116 (27)	3.8 (3.8)	20 (16)	0.7 (0.9)
Aspectos formales	71 (17)	2.4 (3.9)	20 (16)	0.7 (1.5)
Adecuación	61 (14)	2 (2.8)	21 (17)	0.7 (1.8)
Cohesión	63 (15)	2.1 (2.9)	19 (15)	0.6 (1.2)
Intertextualidad	51 (12)	1.7 (2.6)	28 (22)	0.9 (1.3)
Ortografía	54 (13)	1.8 (2.6)	6 (5)	0.2 (0.5)
Coherencia	8 (2)	0.3 (0.5)	13 (10)	0.4 (1)
Total	424 (100)	14.1 (13)	127 (100)	4.2 (5.3)

Fuente: elaboración propia.

con respecto al promedio observamos que es mayor entre los comentarios incorporados, especialmente en el caso de la puntuación y los aspectos formales.

De los comentarios que sí se incorporaron a la versión final del trabajo académico, la mayoría son correctos y, en consecuencia, mejoran el trabajo. Parte del alumnado, en cambio, reflexiona a partir del comentario y empeora la versión final. Además, hay un tercer grupo que, a raíz de la sugerencia de revisión, llega incluso a mejorar la propia propuesta recibida durante la práctica de evaluación. A continuación, en la Tabla 8, mostramos ejemplos de estas tres formas de incorporar los comentarios a la versión final.

Tabla 8. Resultado de la modificación: contraste entre el borrador y la versión final

Primera versión	Propuesta de revisión	Versión final	Resultado tras la modificación
<i>Además es necesario que...</i>	Puntuación: además, es necesario (después de este tipo de marcadores va coma)	<i>Además, es necesario que...</i>	Mejora el trabajo porque incorpora la propuesta recibida
<i>La realidad es que, para el niño, el juego es su forma de vivir</i>	Puntuación: la realidad para el niño,	<i>En la realidad del niño, el juego es su forma de vivir</i>	Mejora el trabajo porque reflexiona y mejora la propuesta recibida
<i>Una vez vista la perspectiva del profesorado de cada una de las modalidades, con sus pros y contras, los autores (Videl y Grau Rubio) llegan a la conclusión de que la modalidad más adecuada para...</i>	Puntuación: Pondría una coma [después de conclusión] ya que procede a dar una explicación.	<i>Una vez visto el punto de vista del personal docente de cada una de las modalidades, con sus pros y contras, se llega a la conclusión que la modalidad más adecuada para...</i>	Empeora el trabajo tras la propuesta recibida

Fuente: elaboración propia.

En cuanto a si la propuesta incorporada es correcta, la gran mayoría de los comentarios incorporados a la versión final son correctos (81.8 por ciento) o de estilo (9.2 por ciento) frente a los incorrectos (9 por ciento), tal y como se puede ver en la Tabla 9. No obstante, llama la atención que, de los no incorporados, la mayoría son correctos (53.5 por ciento) o de estilo (12.6 por ciento), aunque también hay un porcentaje nada despreciable de comentarios incorrectos (33.9 por ciento). En definitiva, hay un número llamativo de comentarios correctos (68) que no fueron incorporados a los trabajos y de comentarios incorrectos (38) que sí figuran en la versión final.

Tabla 9. Tipos de comentarios incorporados y no incorporados

Tipo de comentario	Incorporado		No incorporado	
	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)
Correcto	347 (81.8)	11.5 (10.5)	68 (53.5)	2.3 (3.6)
Incorrecto	38 (9)	1.3 (1.8)	43 (33.9)	1.4 (1.8)
Estilo	39 (9.2)	1.3 (2.5)	16 (12.6)	0.5 (1.3)
Total	424 (100)		127 (100)	

Fuente: elaboración propia.

En lo que respecta a si la propuesta incorporada está formulada directa o indirectamente, la Tabla 10 muestra que los comentarios directos se incorporan en mayor medida que los indirectos, independientemente de que sean correctos (231 frente a 151), incorrectos (31 frente a 11) o estilísticos (29 frente a 10). Esto es especialmente llamativo en el caso de los comentarios correctos, pues se incorpora 90.2 por ciento de los que están formulados directamente y 75.9 por ciento de los que se formulan de forma indirecta. Por lo tanto, hay un 24.1 por ciento de comentarios correctos que probablemente no fueron incorporados porque, al estar formulados indirectamente, exigen un mayor esfuerzo cognitivo por parte del alumnado.

Tabla 10. Tipos de comentarios directos e indirectos incorporados

Tipo de comentario	Incorporado		No incorporado		Total
	Núm. (%)	\bar{X} (σ)	Núm. (%)	\bar{X} (σ)	
Correcto directo	231 (90.2)	8 (8.9)	25 (9.8)	1 (1.6)	256 (100)
Correcto indirecto	151 (75.9)	5 (4.5)	48 (24.1)	2 (3.1)	199 (100)
Incorrecto directo	31 (60.8)	1 (1.7)	20 (39.2)	1 (1)	51 (100)
Incorrecto indirecto	11 (28.2)	0 (0.6)	28 (71.8)	1 (1.4)	39 (100)
Estilístico directo	29 (72.5)	1 (2.5)	11 (27.5)	0 (0.9)	40 (100)
Estilístico indirecto	10 (55.6)	0 (0.5)	8 (44.4)	0 (0.8)	18 (100)
Total	463		140		

Fuente: elaboración propia.

En cuanto a las medidas de dispersión, cabe destacar que el valor de la desviación estándar es mayor en los comentarios incorporados correctos directos (8.9) e indirectos (4.5), y los no incorporados correctos indirectos (3.1), por lo que hay grandes diferencias entre quienes incorporan numerosos de estos comentarios y quienes apenas los incluyen en la versión final.

INTERPRETACIÓN Y DISCUSIÓN DE RESULTADOS

En cuanto al primer objetivo de investigación, esto es, identificar qué tipo de errores corrige el alumnado al evaluar la monografía de una compañera o un compañero, se encontró un gran número de errores de puntuación, seguidos de comentarios en torno a aspectos formales del trabajo y, en menor medida, a cuestiones de cohesión, intertextualidad y ortografía. Los problemas de coherencia textual son los menos corregidos. Así pues, en general, hay una mayor atención hacia los aspectos superficiales de la escritura. La falta de conocimiento gramatical y las dificultades para utilizar de forma instrumental este conocimiento a la hora de revisar los textos condiciona el proceso de escritura y también de revisión de textos (Schere, 2020; Rodríguez Gonzalo, 2012). La mayor corrección de errores de puntuación puede deberse a que se dedicó una sesión y varias actividades a la utilización correcta de los signos de puntuación. De hecho, en otros estudios se ha constatado que el contenido de la intervención didáctica puede condicionar el foco de la revisión de los textos (López-Flamarique *et al.*, 2023). Asimismo, hay una gran variabilidad en el número de correcciones realizadas por cada estudiante. Una de las explicaciones posibles sería la diferente motivación del alumnado a la hora de realizar esta actividad y su implicación en la misma (Yu, 2021). No obstante, también puede estar relacionado con diferencias de nivel en la competencia escrita del alumnado y el tipo de emparejamiento: si a un estudiante con un bajo nivel de competencia escrita le toca revisar un texto escrito que no tiene muchos problemas de escritura, es probable que le cueste detectar los aspectos a ser corregidos. De hecho, se ha visto que el tipo de emparejamiento incide en los resultados de la corrección (Allen y Mills, 2016).

En general, se puede señalar que los resultados de este estudio están en línea con los de Alshuraidah y Storch (2019), quienes encontraron que la mayor parte de comentarios se referían a aspectos superficiales de la lengua, especialmente a cuestiones de puntuación y ortografía. La literatura subraya la necesidad de enseñar y entrenar la práctica de evaluación entre pares, con el objetivo de que el alumnado tome confianza en su rol como revisor y no se limite a corregir aspectos superficiales de la escritura (Baker, 2016; Yu, 2021; Zhan, 2024). En la misma línea, Yu (2021) observó que más de la mitad del *feedback* entre estudiantes se refería a aspectos lingüísticos. En contraposición,

Castelló *et al.* (2010) y Yu (2021) observaron que los aspectos sintácticos y de cohesión textual eran los más comentados.

En este estudio, el número de comentarios referidos a aspectos intertextuales y de coherencia textual, relacionados con las características del género discursivo, recibieron menor número de correcciones. Cabe señalar que estos aspectos eran los más novedosos para el alumnado, pues era la primera vez que se enfrentaban a la utilización de las normas APA y las características de un texto académico. En este sentido, Yu (2021) encontró que una de las dificultades del alumnado al proporcionar *feedback* relacionado con el texto académico es el desconocimiento de las exigencias de este tipo de género. Asimismo, al tratarse de estudiantes del primer semestre del grado, sus conocimientos sobre los temas abordados en las monografías eran limitados, lo que dificultaba realizar comentarios acerca del contenido del texto. Por último, la falta de atención a aspectos como la ortografía —encontrada en esta investigación— se puede deber a que el propio programa ofimático realiza la corrección de la mayor parte de errores ortográficos de los textos.

Al clasificar el tipo de comentarios de corrección realizados, encontramos que los errores relacionados con la puntuación normalmente se clasifican de forma correcta, algo que no ocurre con la misma frecuencia al hablar de la cohesión y la adecuación. Esto puede deberse a que el primer tipo de error se localiza más fácilmente en la superficie del texto. En cambio, identificar errores de cohesión entraña un mayor dominio de la competencia lingüística y metalingüística (Pitt *et al.*, 2020; Hill y West, 2020, cit. por Cano *et al.*, 2022).

En cuanto al tipo de comentario, directo o indirecto, a pesar de que la instrucción docente era que se diera una explicación razonada de las correcciones (observación indirecta), la mayor parte de los comentarios son directos, pues exigen un menor manejo de nociones gramaticales. De hecho, en muchas ocasiones, la revisión requiere de conocimientos que no están explícitos en la conciencia metalingüística del alumnado. Tal y como se ha visto en este estudio, algunas veces es capaz de detectar inadecuaciones textuales, pero no puede explicar las causas ni resolver el problema (Rodríguez Gonzalo, 2012).

En cuanto al segundo objetivo de investigación, cabe hacer hincapié en que la mayoría de los comentarios realizados durante la práctica de evaluación entre iguales son correctos. Estos resultados están en sintonía con otros estudios (Yu, 2021). Por eso, coincidimos en destacar este tipo de evaluación como una herramienta de aprendizaje innovadora y con enorme potencial en las aulas de educación superior. Con todo, hemos comprobado que, en ocasiones, las propuestas realizadas pueden ser incorrectas, algo que no necesariamente repercute en la calidad del texto final. De hecho, hemos constatado cómo algunos de los comentarios incorrectos no se incorporan a la última versión. De todas formas, consideramos que estas observaciones también resultan beneficiosas durante el proceso de enseñanza-aprendizaje de la escritura académica.

mica, pues fomentan la reflexión y la actividad metalingüística del alumnado que debe decidir si incorporar o no una determinada alternativa (Dressler *et al.*, 2019). En este sentido, vale la pena mencionar que una gran parte de los comentarios incorrectos se referían a cuestiones de intertextualidad y coherencia relacionados con el género discursivo, los cuales, como ya se ha mencionado, eran los más novedosos para el alumnado. Por todo ello, y al tratarse de alumnado de primer curso, consideramos que seguirán adquiriendo estos conocimientos a lo largo del grado.

Por último, en lo que concierne al tercer objetivo planteado, constatamos que 81.8 por ciento de los comentarios correctos, recibidos durante la práctica de evaluación entre iguales, se incorporan a la versión final del trabajo académico. Por lo tanto, estamos de acuerdo con otros autores en que la evaluación entre iguales conlleva una mejora de la producción escrita del alumnado (Cho y Cho, 2011; Dressler *et al.*, 2019). Cabe notar que cuando el *feedback* proporcionado es directo, la probabilidad de que el comentario se incorpore al texto final es mayor.

CONCLUSIONES

En resumen, tanto el número y tipo de comentarios realizados como su correcta clasificación, sumado al hecho de que la mayor parte de las observaciones realizadas son acertadas, deja claro que esta práctica de evaluación promueve la reflexión metalingüística y metacognitiva del alumnado (Boud, 1995; Rijlaarsam y Couzijn, 1999). Esto, sin duda, contribuye al desarrollo de su competencia escrita, tal y como se ha comprobado en estudios previos (Paulus, 1999; Wu y Schunn, 2021), y a la mejora de su producción escrita (Huisman *et al.*, 2019). Sin embargo, se ha observado que en muchas ocasiones el alumnado presenta dificultades para razonar sus propuestas de corrección. En este sentido, subrayamos la importancia de incorporar la reflexión metalingüística en las diferentes fases del proceso de escritura y avanzar hacia una enseñanza de la gramática que permita al alumnado poner en relación los conocimientos gramaticales con la elaboración de textos orales y escritos (Camps y Ribas, 2017; Rodríguez Gonzalo, 2012).

Del mismo modo, se detectó la necesidad de trabajar en mayor profundidad aspectos relacionados con la coherencia, pues es lo que más le cuesta corregir al alumnado. En este sentido, desde una perspectiva sociocultural, sería interesante vincular la monografía evaluada con alguna otra asignatura del primer curso. De esta forma, se podría trabajar el contenido con mayor profundidad y presentar el resultado final de estas producciones escritas ante la comunidad académica con el objetivo de fomentar la motivación, pertenencia y responsabilidad ante el trabajo realizado.

Asimismo, tras la aparición de plataformas de inteligencia artificial generativa como ChatGPT o Gemini, cuyo uso se ha extendido de forma exponencial entre el alumnado universitario, es preciso repensar las prioridades de este proceso de alfabetización y las nuevas estrategias metalingüísticas y metacognitivas que entran en juego en el proceso de escritura. Algunas experiencias, por ejemplo, han comenzado a comparar la evaluación entre pares con la evaluación de textos realizada por ChatGPT (Kalantzis y Cope, 2024).

Este trabajo es un estudio de caso de carácter exploratorio y los resultados deben ser confirmados con nuevas investigaciones. Además, debido al tamaño de la muestra, la validez externa es limitada, por lo que los resultados no pueden generalizarse; no obstante, sí presentan una gran validez ecológica, pues se trabajó con un grupo intacto y la práctica analizada está inmersa en el currículo de una asignatura obligatoria del grado. De cara a futuras investigaciones, como respuesta a la gran dispersión de datos encontrada, sería interesante tener en cuenta el nivel de competencia escrita del alumnado al formar las parejas para la práctica de evaluación y contrastar la revisión del alumnado con la realizada posteriormente por la docente, esto con el fin de detectar posibles errores que no hayan sido identificados en la evaluación entre pares. Además, sería pertinente aumentar el tamaño de la muestra y analizar los resultados desde una perspectiva de género para identificar si el hecho de ser hombre o mujer influye en el tipo de *feedback* proporcionado.

Por último, a pesar de los numerosos beneficios destacados a lo largo del trabajo sobre la evaluación entre iguales, se ha comprobado que esta práctica resulta exigente para el alumnado. Por lo tanto, si bien es una herramienta de gran utilidad didáctica, requiere que en las aulas se fomente la reflexión metalingüística del alumnado y que reciba formación específica y herramientas para realizar este tipo de evaluación. Esperamos que este tipo de estudios sirvan para romper con las siempre resistentes inercias en las prácticas educativas y contribuya a incorporar la evaluación entre iguales como una herramienta didáctica que facilita el desarrollo de la competencia escrita y metacognitiva del alumnado.

REFERENCIAS

- ALLEN, David y Amy Mills (2016), “The Impact of Second Language Proficiency in Dyadic Peer Feedback”, *Language Teaching Research*, vol. 20, núm. 4, pp. 498-513. DOI: <https://doi.org/10.1177/1362168814561902><https://doi.org/10.1177/1362168814561902>
- ALSHURAIDAH, Ali y Noemi Storch (2019), “Investigating a Collaborative Approach to Peer Feedback”, *ELT Journal*, vol. 73, núm. 2, pp. 166-174. DOI: <https://doi.org/10.1093/elt/ccy057>

- AZNÁREZ-Mauleón, Mónica, Maite López-Flamarique e Isabel García-Del-Real (2020), “La actividad metalingüística en una tarea de escritura colaborativa en educación primaria”, *Ocnos. Revista de Estudios sobre la Lectura*, vol. 19, núm. 3, pp. 42-54. DOI: https://doi.org/10.18239/OCNOS_2020.19.3.2265
- BAJTÍN, Mijaíl (1992), *Estética da criação verbal*, São Paulo, Martins Fontes.
- BAKER, Kimberly M. (2016), “Peer Review as a Strategy for Improving Students’ Writing Process”, *Active Learning in Higher Education*, vol. 17, núm. 3, pp. 179-192. DOI: <https://doi.org/10.1177/1469787416654794>
- BOGDAN, Robert C. y Sari Knopp Biklen (1997), *Qualitative Research for Education: An introduction to theory and methods*, Massachusetts, Allyn & Bacon.
- BOUD, David (1990), “Assessment and the Promotion of Academic Values”, *Studies in Higher Education*, vol. 15, núm. 1, pp. 101-111. DOI: <https://doi.org/10.1080/03075079012331377621>
- BOUD, David (1995), “Assessment and Learning: Contradictory or complementary”, en Peter Knight (ed.), *Assessment for Learning in Higher Education*, Londres, Kogan, pp. 35-48.
- BUENO-Alastuey, María Camino, Raychel Vasseur e Idoia Elola (2022), “Effects of Collaborative Writing and Peer Feedback on Spanish as a Foreign Language Writing Performance”, *Foreign Language Annals*, vol. 55, núm. 2, pp. 517-539. DOI: <https://doi.org/10.1111/flan.12611>
- CAMPS Mundó, Anna y Montserrat Castelló Badía (2013), “La escritura académica en la universidad”, *Revista de Docencia Universitaria*, vol. 11, núm. 1, pp. 17-36.
- CAMPS, Anna y Marta Milian (2000), “La actividad metalingüística en el aprendizaje de la escritura”, en Anna Camps y Marta Milian (eds.), *El papel de la actividad metalingüística en el aprendizaje de la escritura*, Rosario, Homo Sapiens Ediciones, pp. 7-39.
- CAMPS, Anna y Teresa Ribas (coords.) (2017), *El verbo y su enseñanza: hacia un modelo de enseñanza de la gramática basado en la actividad reflexiva*, Barcelona, Octaedro.
- CANO García, Elena, Laura Pons-Seguí y Maite Fernández-Ferrer (2022), “Efectos de los tipos de *feedback* entre iguales en los trabajos escritos”, *Profesorado. Revista de Currículum y Formación del Profesorado*, vol. 26, núm. 1, pp. 127-148. DOI: <https://doi.org/10.30827/profesorado.v26i1.13929>
- CARLINO, Paula (2005), *Escribir, leer y aprender en la universidad. Una introducción a la alfabetización académica*, México, Fondo de Cultura Económica.
- CARLINO, Paula, Patricia Iglesia e Irene Laxalt (2013), “Concepciones y prácticas declaradas de profesores terciarios en torno al leer y escribir en las asignaturas”, *Revista de Docencia Universitaria*, vol. 11, núm. 1, pp. 105-136.
- CASSANY, Daniel (1993), *Reparar la escritura. Didáctica de la corrección de lo escrito*, Barcelona, Graó.
- CASSANY, Daniel (1999), *Construir la escritura*, Barcelona, Paidós.

- CASTELLÓ, Manuel (2007), “El proceso de composición de textos académicos”, en Montserrat Castelló Badia (coord.), *Escribir y comunicarse en contextos científicos y académicos: conocimientos y estrategias*, Barcelona, Graó, pp. 47-82.
- CASTELLÓ, Montserrat (2000), “Las concepciones de los estudiantes sobre la escritura académica”, en Anna Camps y Marta Milian (eds.), *El papel de la actividad metalingüística en el aprendizaje de la escritura*, Rosario, Homo Sapiens Ediciones, pp. 67-104.
- CASTELLÓ, Montserrat, Gerardo Bañales Faz y Norma Alicia Vega López (2010a), “Enfoques en la investigación de la regulación de escritura académica: estado de la cuestión”, *Electronic Journal of Research in Educational Psychology*, vol. 8, núm. 22, pp. 1253-1282.
- CASTELLÓ, Montserrat, Dolores González y Ana Iñesta (2010b), “La regulación de la escritura académica en el doctorado: el impacto de la revisión colaborativa en los textos”, *Revista Española de Pedagogía*, vol. 68, núm. 247, pp. 521-537, en: <https://www.jstor.org/stable/23766358> (consulta: 26 de noviembre de 2023).
- CHO, Young Ho y Kwangsu Cho (2011), “Peer Reviewers Learn from Giving Comments”, *Instructional Science*, vol. 39, núm. 5, pp. 629-643. DOI: <https://doi.org/10.1007/s11251-010-9146-1>
- DOMINICK Peter G., Richard R. Reilly y Jack W. McGourty (1997), “The Effects of Peer Feedback on Team Member Behavior”, *Group & Organization Management*, vol. 22, núm. 4, pp. 508-520. DOI: <https://doi.org/10.1177/1059601197224006>
- DRESSLER, Roswita, Mon-Wai Chu, Katie Crossman y Brianna Hilman (2019), “Quantity and Quality of Uptake: Examining surface and meaning-level feedback provided by peers and an instructor in a graduate research course”, *Assessing Writing*, vol. 39, pp. 14-24. DOI: <https://doi.org/10.1016/j.asw.2018.11.001>
- FLOWER, Linda y John R. Hayes (1981), “A Cognitive Process Theory of Writing”, *College Composition and Communication*, vol. 32, núm. 4, pp. 365-387. DOI: <https://doi.org/10.2307/356600>
- GOMBERT, Jean Émile (1992), *Metalinguistic Development*, Chicago, The University of Chicago Press.
- GÓMEZ, Miguel Ángel y Victoria Quesada Serra (2017), “Coevaluación o evaluación compartida en el contexto universitario: la percepción del alumnado de primer curso”, *Revista Iberoamericana de Evaluación Educativa*, vol. 10, núm. 2, pp. 9-30. DOI: <https://doi.org/10.15366/rie2017.10.2.001>
- GÓMEZ-RUIZ, Miguel Ángel, Gregorio Rodríguez-Gómez y María Soledad Ibarra-Sáiz (2013), “COMPES. Autoinforme sobre competencias básicas relacionadas con la evaluación de los estudiantes universitarios”, *Estudios sobre Educación*, vol. 24, pp. 197-224. DOI: <https://10.15581/004.24.2031>
- HADWIN, Allyson F., Mika Oshige, Carmen L.Z. Gress y Philip H. Winne (2010), “Innovative Ways for Using gStudy to Orchestrate and Research Social Aspects of Self-Regulated Learning”, *Computers in Human Behavior*, vol. 26, núm. 5, pp. 794-805.

- HUGO Rojas, Evelyn Susana y Alejandra Cecilia Meneses Arévalo (2021), “Reflexión metalingüística en L1 y escritura. Revisión de investigaciones en contextos escolares”, *Literatura y Lingüística*, vol. 44, pp. 315-344. DOI: <https://doi.org/10.29344/0717621X.44.2339>
- HUISMAN, Bart, Nadira Saab, Paul van den Broek y Jan van Driel (2019), “The Impact of Formative Peer Feedback on Higher Education Students’ Academic Writing: A meta-analysis”, *Assessment & Evaluation in Higher Education*, vol. 44, núm. 6, pp. 863-880. DOI: <https://doi.org/10.1080/02602938.2018.1545896>
- HYLAND, Ken (2004), *Genre and Second Language Writing*, Michigan, The University of Michigan Press.
- IBARRA-Sáiz, María Soledad, Gregorio Rodríguez-Gómez, David Boud, Tijs Rotsaert, Sally Brown, Marta Lorena Salinas Salazar e Hilda Mar Rodríguez Gómez (2020), “El futuro de la evaluación en la educación superior”, *Relieve. Revista Electrónica de Investigación y Evaluación Educativa*, vol. 26, núm. 1, pp. 1-6. DOI: <https://doi.org/10.7203/relieve.26.1.17323>
- KALANTZIS, Mary y Bill Cope (2024), “Literacy in the Time of Artificial Intelligence”, *Reading Research Quarterly*, vol. 60, núm. 1. DOI: <https://doi.org/10.1002/rrq.591>
- LEA, Mary R. y Brian V. Street (1998), “Student Writing in Higher Education: An academic literacies approach”, *Studies in Higher Education*, vol. 23, núm. 2, pp. 157-72. DOI: <https://doi.org/10.1080/03075079812331380364>
- LIU, Ngai-Fun y David Carless (2006), “Peer Feedback: The learning element of peer assessment”, *Teaching in Higher Education*, vol. 11, núm. 3, pp. 279-290. DOI: <https://doi.org/10.1080/13562510600680582>
- LOBATO Fraile, Clemente (2006), “Estudio y trabajo autónomos del estudiante”, en Miguel Díaz (ed.), *Metodologías de enseñanza y aprendizaje para el desarrollo de competencias: orientación para el profesorado universitario ante el espacio europeo de educación*, Madrid, Alianza Editorial, pp. 191-224.
- LÓPEZ-Flamarique, Maite, Mónica Aznárez-Mauleón e Isabel García-del-Real (2023), “¿Cómo planifica y revisa el alumnado de primaria? Un estudio de caso de una tarea de escritura colaborativa”, *Didáctica. Lengua y Literatura*, vol. 35, pp. 113-125. DOI: <https://doi.org/10.5209/dill.81810>
- LUNDSTROM, Kristi y Wendy Baker (2009), “To Give is Better than to Receive: The benefits of peer review to the reviewer’s own writing”, *Journal of Second Language Writing*, vol. 18, núm. 1, pp. 30-43. DOI: <https://doi.org/10.1016/j.jslw.2008.06.002>
- NAVARRO, Federico (2018), “Didáctica basada en géneros discursivos para la lectura, escritura y oralidad académicas”, en Federico Navarro y Graciela Aparicio (coords.), *Manual de lectura, escritura y oralidad académicas para ingresantes a la Universidad*, Bernal, Universidad Nacional de Quilmes, pp. 13-25.

- NICOL, David y Suzanne McCallum (2022), “Making Internal Feedback Explicit: Exploiting the multiple comparisons that occur during peer review”, *Assessment & Evaluation in Higher Education*, vol. 47, núm. 3, pp. 424-443. DOI: <https://doi.org/10.1080/02602938.2021.1924620>
- PANADERO, Ernesto, Gavin T. L. Brown y Jan-Willem Strijbos (2016), “The Future of Student Self-Assessment: A review of known unknowns and potential directions”, *Educational Psychology Review*, vol. 28, pp. 803-830. DOI: <https://doi.org/10.1007/s10648-015-9350-2>
- PANADERO, Ernesto, Juan Fraile, Javier Fernández Ruiz, David Castilla-Estévez y Miguel A. Ruiz (2019), “Spanish University Assessment Practices: Examination tradition with diversity by faculty”, *Assessment & Evaluation in Higher Education*, vol. 44, núm. 3, pp. 379-397. DOI: <https://doi.org/10.1080/02602938.2018.1512553>
- PAULUS, Trena M. (1999), “The Effect of Peer and Teacher Feedback on Student Writing”, *Journal of Second Language Writing*, vol. 8, núm. 3, pp. 265-289. DOI: [https://doi.org/10.1016/S1060-3743\(99\)80117-9](https://doi.org/10.1016/S1060-3743(99)80117-9)
- PITT, Edd, Margaret Bearman y Rachele Esterhazy (2020), “The Conundrum of Low Achievement and Feedback for Learning”, *Assessment & Evaluation in Higher Education*, vol. 45, núm. 2, pp. 239-250. DOI: <https://doi.org/10.1080/02602938.2019.1630363>
- REYES, Graciela (1999), *Cómo escribir bien en español*, Madrid, Arco Libros S.L.
- RIJLAARSDAM, Gert y Michel Couzijn (1999), “La estimulación de la metacognición en la enseñanza de la escritura”, en Marta Millian y Anna Camps (comps.), *El papel de la actividad metalingüística en el aprendizaje de la escritura*, Rosario, Homo Sapiens, pp. 215-260.
- RODRÍGUEZ Gómez, Gregorio, María Soledad Ibarra y Eduardo García Jiménez (2013), “Autoevaluación, evaluación entre iguales y coevaluación: conceptualización y práctica en las universidades españolas”, *Revista de Investigación en Educación*, vol. 11, núm. 2, pp. 198-210.
- RODRÍGUEZ Gómez, Gregorio, María Soledad Ibarra y Miguel Ángel Gómez Ruiz (2011), “e-Autoevaluación en la universidad: un reto para profesores y estudiantes”, *Revista de Educación*, vol. 356, pp. 401-430.
- RODRÍGUEZ Gonzalo, Carmen (2012), “La enseñanza de la gramática: las relaciones entre la reflexión y el uso lingüístico”, *Revista Iberoamericana de Educación (RIE)*, vol. 59, pp. 87-118. DOI: <https://doi.org/10.35362/rie590458>
- SANMARTI Puig, Neus (2007), *10 ideas clave: evaluar para aprender*, Barcelona, Graó.
- SCARDAMALIA, Marlene y Carl Bereiter (1992), “Dos modelos explicativos de los procesos de composición escrita”, *Infancia y Aprendizaje*, vol. 15, núm. 58, pp. 43-64. DOI: <https://doi.org/10.1080/02103702.1992.10822332>
- SCHERE, Jimena (2020), “Escritura académica y reflexión gramatical en el comienzo de la formación universitaria”, *Álabe*, vol. 22, DOI: <https://doi.org/10.15645/Alabe2020.22.7>

- SCHNEUWLY, Bernard (2007), “Genres écrits et oraux et forme scolaire: enseignement et apprentissage de la langue première à l’école”, en Catherine Boré (ed.), *Construire et exploiter des corpus de genres scolaires*, Namur, Presses Universitaires de Namur, pp. 13-26.
- SCRIVEN, Michael (1967), “The Methodology of Evaluation”, en Ralph Winfred Tyler, Robert M. Gagne y Michael Scriven (eds.), *Perspectives of Curriculum Evaluation*, Quebec, Rand McNally, pp. 39-83.
- SIMONS, Helen (2011), *El estudio de caso: teoría y práctica*, Madrid, Morata.
- STAKE, Robert E. (1999), *Investigación con estudio de caso*, Madrid, Morata.
- TOLCHINSKY, Liliana (2000), “Las concepciones de los estudiantes sobre la escritura académica”, en Anna Camps y Marta Milian (eds.), *El papel de la actividad metalingüística en el aprendizaje de la escritura*, Rosario, Homo Sapiens Ediciones, pp. 39-67.
- TOPPING, Keith (1998), “Peer Assessment between Students in Colleges and Universities”, *Review of Educational Research*, vol. 68, núm. 3, pp. 249-276. DOI: <https://doi.org/10.3102/00346543068003249>
- WU, Yon y Christian D. Schunn (2021), “The Effects of Providing and Receiving Peer Feedback on Writing Performance and Learning of Secondary School Students”, *American Educational Research Journal*, vol. 58, núm. 3, pp. 492-526. DOI: <https://doi.org/10.3102/0002831220945266>
- YARROW, Fiona y Keith J. Topping (2001), “Collaborative Writing: The effects of metacognitive prompting and structured peer interaction”, *British Journal of Educational Psychology*, vol. 71, núm. 2, pp. 261-282. DOI: <https://doi.org/10.1348/000709901158514>
- YAZAN, Bedrettin (2015), “Three Approaches to Case Study Methods in Education: Yin, Merriam, and Stake”, *The Qualitative Report*, vol. 20, núm. 2, pp. 134-152, en: <http://nsuworks.nova.edu/tqr/vol20/iss2/12> (consulta: 12 de febrero de 2024).
- YIN, Robert K. (2009), *Case Study Research. Design and methods*, California, Sage. DOI: <https://doi.org/10.33524/cjar.v14i1.73>
- YU, Shulin (2021), “Giving Genre-Based Peer Feedback in Academic Writing: Sources of knowledge and skills, difficulties and challenges”, *Assessment and Evaluation in Higher Education*, vol. 46, núm. 1, pp. 36-53. DOI: <https://doi.org/10.1080/02602938.2020.1742872>
- ZHAN, Ying (2024), “Are they Ready? An investigation of university students’ difficulties in peer assessment from dual perspectives”, *Teaching in Higher Education*, vol. 29, núm. 4, pp. 823-840. DOI: <https://doi.org/10.1080/13562517.2021.2021393>
- ZIMMERMAN, Barry J. (2000), “Attainment of Self-Regulation: A social cognitive perspective”, en Monique Boekaerts, Paul R. Pintrich y Moshe Zeidner (eds.), *Handbook of Self-Regulation, Research, and Applications*, Orlando, Academic Press, pp. 13-39.