ANALES DE FILOLOGÍA FRANCESA

Análisis de los errores formales, gramaticales y sintácticos de los estudiantes universitarios de francés

An analysis of the formal, grammatical and syntactical errors of university students of French

OLGA ROCÍO SERRANO-CARRANZA
Universidad Nacional de Colombia
orserranoc@unal.edu.co

MARTHA LEONOR MARTÍNEZ ACOSTA
Universidad Nacional de Colombia
mlmartineza@unal.edu.co

Recibido: 12-08-2025 | Aceptado: 12-09-2026

Cómo citar este artículo: Serrano-Carranza, Olga Rocío, & Martínez Acosta, Martha Leonor. (2026). Análisis de los errores formales, gramaticales y sintácticos de los estudiantes universitarios de francés. *Anales de Filología Francesa*, 34, pp. 211-231. https://doi.org/10.6018/analesff.675691

Abstract:
This investigation was based on an analysis of the most common formal, grammatical and syntactic errors in an annotated corpus of compositions (22,791 words) written by 46 university students of French. It compared their texts with the information in a survey of how 23 of the students perceived their mistakes. In total, 4,131 mistakes were found in their writings. The most frequent mistakes, in descending order, were to do with spelling (31.3%), punctuation (23.3%) and the correct use of verbs (10.3%), prepositions (7.4%) and articles (6.9%). According to the results of the survey of perceptions, a correct use of spelling was also thought to be the most complicated problem (73.9%). Thanks to this study, it was possible to spot the most common mistakes, with the aim of devising long-term metacognitive and personalized corrective measures and thus enable the students to perfect their writing skills.

Key words:
Mistakes, learner corpus, annotated corpus, writing skills, studying French as a foreign language.

Résumé:
Cette recherche s'est fondée sur l'étude des erreurs formelles, grammaticales et syntaxiques, les plus fréquentes, qui ont été repérées dans un corpus annoté de productions écrites (22 791 mots) de 46 étudiants universitaires de français. Une comparaison a aussi été établie entre ces données et les résultats d'une enquête menée auprès de 23 étudiants sur leurs fautes linguistiques.
Au total, 4 131 erreurs ont été trouvées dans le corpus. Les erreurs les plus courantes sont, en ordre décroissant, les orthographiques (31,3 %), celles en rapport avec les signes de ponctuation (23,3 %), les verbes (10,3 %), les prépositions (7,4 %) et les articles (6,9 %). Les erreurs orthographiques ont également été cataloguées comme étant les plus difficiles pour les étudiants interrogés (73,9 %). L'identification de ces erreurs permettra de proposer, à long terme, des activités métacognitives et correctives pour que les apprenants améliorent leur écriture.

Mots-clés:
Erreur, corpus d'apprenants, corpus annoté, compétence écrite, apprentissage du Français Langue Étrangère.

1. Introducción

La destreza en la escritura comprende el desarrollo de la "competencia lingüística, discursiva, sociolingüística y estratégica" (Madrid & Corral, 2018, p. 182) que el estudiante aplica para "expresar ideas, pensamientos, información y sentimientos de manera clara, coherente y efectiva a través de textos escritos" (Moreira & Montero, 2025, p. 108). En el contexto de aprendizaje de la lengua extranjera, el proceso de producción escrita involucra una o varias etapas entre las cuales se mencionan primordialmente la de preescritura, escritura, corrección y la de presentación del texto final; su presencia depende de la intencionalidad formal o informal del texto, que se singulariza además, por una tonalidad, un registro, un vocabulario, estructuras lingüísticas y un estilo de escritura particular (Universidad de Antioquia, 2025).

La etapa de la preescritura comprende la generación de ideas sobre las temáticas de escritura y algunas precisiones sobre el tipo de texto. Asimismo, "necesita aclarar los fines y al lector a quien se dirige" (Agurto et al., 2019, p. 24). Para facilitar los mecanismos creativos, es posible hacer uso de "la técnica de la lluvia de ideas, la cual lleva a pensar sin inhibición que la finalidad es producir tantas ideas como sea posible sobre un determinado tópico o problema" (Rodríguez, 2022, p. 109). Otros métodos son el de la lista de ideas, la creación de mapas mentales, de esquemas, planes o bosquejos en los que se resaltan los pensamientos que van a ser desarrollados, organizados y entrelazados (Rodríguez, 2022). También es clave contar con un bagaje de conocimientos previos (incluso en lengua materna y en una o más lenguas extranjeras) que actúan como facilitadores en los procesos reflexivos y que se enriquecen con la lectura de diversas fuentes bibliográficas.

Tras la fase de preescritura, vienen luego la de escritura y la de corrección del texto. En estas, se plasman y se profundizan las ideas por medio de oraciones y párrafos; se presta especial atención a la claridad, a la coherencia, a la cohesión, al empleo de estructuras lingüísticas y sociodiscursivas de la lengua extranjera, a la riqueza del vocabulario y a su uso adecuado. De modo similar, se toma en consideración la ortografía, los signos de puntuación y el perfeccionamiento en el estilo de redacción (Agurto et al., 2019). Por lo general, la corrección engloba varias revisiones hasta escribir el texto definitivo.

Ahora bien, en el caso de las producciones finales, escritas por usuarios básicos de francés como lengua extranjera, es plausible encontrar aún falencias y, en especial, si las actividades de escritura estaban sujetas a procesos de evaluación con restricciones de tiempo o de uso de materiales de ayuda como diccionarios, traductores, libros de gramática o herramientas de aprendizaje virtual.

Con el ánimo de identificar, clasificar y estudiar los errores en las producciones escritas de un grupo de estudiantes de francés, en contextos evaluativos, se emprendió esta investigación. Las preguntas a las que se busca responder son las siguientes: ¿Cuáles son los errores más prevalentes en un corpus de interlengua de estudiantes universitarios de francés en estadios iniciales? ¿Cuáles son las creencias de los estudiantes sobre los errores más frecuentes que cometen en sus producciones escritas al momento de responder a la encuesta que les fue suministrada? Como respuesta a estos interrogantes se plantea que el mayor número de errores se presentará a nivel ortográfico como consecuencia de la marcada opacidad de la lengua francesa.

2. Marco teórico

En esta sección, se definen las nociones de lengua opaca y error. Se detalla la clasificación de los errores. Del mismo modo, se introducen los conceptos de interlengua,

corpus de interlengua, conciencia metalingüística y su relación con el análisis de errores y el componente pedagógico.

2.1. Lengua opaca, concepto de error y tipología

Un idioma opaco es aquel en el que "the grapheme-phoneme and phoneme-grapheme correspondences are inconsistent" (Dalle, 2024, p. 5). En efecto, en "orthographies commonly described as transparent, spelling can be derived largely from phonological information, whereas in more opaque orthographies, accurate spelling requires the integration of lexical, contextual, and morphological constraints" (Perfetti, 2003; Perfetti & Dunlap, 2008[1]). En este sentido, el carácter opaco de la lengua francesa induce a los estudiantes a cometer errores en sus producciones escritas.

Los errores son definidos como "desviaciones sistemáticas de las normas de la lengua objetivo, originadas porque el aprendiz aún no ha internalizado la forma correcta o no conoce las reglas" (Corder, 1975; Ellis et al., 2020; Larsen, 1991[2]). Los errores han sido catalogados bajo ciertos parámetros como lo son "el lingüístico, el etiológico, el pedagógico, el comunicativo, el pragmático" (Vázquez, 2010[3]), entre otros.

Los errores etiológicos se concentran en "las causas de los errores" (Ferreira & Elejalde, 2020, p. 118). Una de ellas consiste en "transférer les règles de la langue maternelle à la langue étrangère" (Afifi, 2018, p. 280). En esta situación particular, el aprendiente usa conocimientos previos que aplica a su actual aprendizaje del idioma extranjero y que también pueden proceder de "otras terceras lenguas aprendidas" (Ferreira & Elejalde, 2020, p. 119). Los errores que resultan de estas influencias son llamados errores interlinguales o interferencias. Otros errores, conocidos como intralinguales, "result from the internal processes of second-language learning. They include rule overgeneralization (*comed, childs*), simplification (*he go home*), and hypercorrection (*she didn't went*)" (Buskivadze & Baratashvili, 2025, p. 349).

Los errores pedagógicos se producen por explicaciones inexactas por parte del profesor; se distinguen por ser temporales, duraderos, individuales o grupales (Vázquez, 2010[4]). Los errores comunicativos emergen por una transmisión defectuosa de los mensajes que conduce a una comprensión o interpretación limitada o insuficiente (García & Zuluaga, 2021). Los errores pragmáticos engloban fallos "discursivos o de pertinencia; reglas relacionadas con la coherencia, la cohesión, la cooperación, los turnos, los silencios y las pausas" (Vázquez, 2010[5]).

Los errores lingüísticos se catalogan, entre otros, en errores formales, gramaticales y sintácticos. Los primeros se refieren a los empleos errados en las palabras de letras, de acentos (errores ortográficos) o de los signos de puntuación "que organizan el discurso para facilitar su comprensión, poniendo de manifiesto las relaciones sintácticas y lógicas entre sus diversos constituyentes, evitando posibles ambigüedades" (Real Academia Española, 2025, párr, 1). Los segundos agrupan las equivocaciones relacionadas con las categorías gramaticales, a saber, determinantes, artículos, pronombres, adjetivos, preposiciones, adverbios, verbos, conjunciones, sustantivos o interjecciones. Los terceros abarcan las omisiones de ciertos elementos en las oraciones, su organización errada o la utilización de una categoría gramatical en vez de otra (Granger et al., 2022).

---

[1] Citado por Comblain (2026, p. 2).
[2] Citado por Macías (2025, p. 316).
[3] Citado por Lara et al. (2023, p. 1828).
[4] Citado por Lara et al. (2023).
[5] Citado por Lara et al. (2023, p. 1829).

2.2. Interlengua y corpus de interlengua

La interlengua es entendida como el "sistema lingüístico producido por un estudiante de una lengua extranjera en cada etapa de su aprendizaje" (Lozano et al., 2025, p. 64). Esta se compone de "diferentes estados que el estudiante atraviesa y de los cuales se apropia para poder expresarse eficientemente" (García & Zuluaga, 2021, p. 348) en la lengua que aprende. Los corpus de interlengua son recopilaciones de "productions orales ou écrites d'apprenants de langues étrangères ou secondes" (Granger, 2018, párr. 4) que "reúnen formas correctas e incorrectas del idioma que el informante está aprendiendo" (Bailini, 2023, p. 19) y cuya finalidad es la de examinar fenómenos lingüísticos, comunicativos o socioculturales, entre los cuales, el análisis de los errores ocupa un interés preponderante (Granger, 2018).

Un corpus de interlengua es transversal cuando contiene informaciones sobre estadios determinados del proceso de aprendizaje, por ejemplo del nivel inicial, intermedio o avanzado. Se hace alusión a un corpus de aprendientes longitudinal cuando las producciones de los hablantes fueron obtenidas en diversos momentos de su aprendizaje; este tipo de corpus permite examinar "le développement langagier des apprenants car il contient les productions d'une cohorte d'apprenants pendant une période de plusieurs mois ou de plusieurs années, ce qui permet de suivre en quelque sorte le processus d'acquisition en cours" (De Cock & Tyne, 2014, párr. 6).

De acuerdo con los objetivos de cada estudio, los corpus se constituyen en función de parámetros particulares. Algunos de ellos toman en consideración ciertas características de los participantes, como, por ejemplo, la edad, el género, la manera en la que aprenden la lengua extranjera, el manejo suplementario de otros idiomas, "el nivel de competencia, la lengua materna y años de estudio de los aprendientes con el objeto de identificar aspectos diferenciadores del proceso de adquisición" (Ferreira & Elejalde, 2020, p. 118). Otros criterios están relacionados con la tarea solicitada para recopilar los datos y con las instrucciones correspondientes: acceso libre o restringido a recursos tales como diccionarios, traductores o sitios de aprendizaje en Internet, límite de tiempo, longitud de la tarea, temáticas, registro de redacción de la tarea, entre otros (Granger et al., 2020).

En lo referente a los corpus de aprendientes escritos, tras haber recolectado las producciones manuscritas de los estudiantes, se procede posteriormente a digitalizarlos, prestando especial atención a no modificarlos. Algunos investigadores trabajan con corpus brutos, "c'est-à-dire sans aucune annotation linguistique" (Granger, 2018, párr. 10). No obstante, antes de efectuar el análisis de un corpus, también es posible marcarlo. Un corpus anotado es aquel al que se le han agregado etiquetas de identificación, acompañadas de descripciones. "Estas marcas se hacen a nivel de palabra, sintagma, oración, fragmento o texto completo" (Aldama et al., 2022, p. 4). Las anotaciones se adicionan manualmente o de manera automática con la ayuda de programas de etiquetado y, de modo general, comprenden algún tipo de "codificación de propiedades lingüísticas relativamente abstractas que no pueden observarse directamente en el texto" (García-Miguel, 2022, p. 18). Las descripciones de las etiquetas incluyen a menudo informaciones relativas a "la morfología, sintaxis, semántica, entidades, correferencias" (Aldama et al., 2022, p. 4), al componente etiológico (Ferreira & Elejalde, 2020), gramatical, ortográfico, comunicativo, sociodiscursivo o sociocultural. Los nombres de las anotaciones son sugeridos por los investigadores, aunque, muchas veces resulte más sencillo, fundamentarse en listas de categorías previamente establecidas por expertos en investigaciones anteriores. Con el propósito de evitar malas categorizaciones, es indispensable corregir varias veces el etiquetado del corpus.

Un corpus anotado facilita el análisis de los datos tanto cuantitativa como cualitativamente. Con la ayuda de herramientas informáticas de análisis textual, es factible

calcular el número de ocurrencias de los elementos del corpus, obtener ejemplos de estos en contextos de uso, compararlos y establecer generalizaciones.

2.3. Conciencia metalingüística y corpus con fines pedagógicos

La conciencia metalingüística se concibe como "la capacidad de reflexión sobre el lenguaje" (Quispe-Morales, 2022, p. 1). Este proceso metacognitivo integra el análisis, el establecimiento de interrelaciones y la selección más acertada entre diversas categorías y estructuras lingüísticas, comunicativas o socioculturales (Ghanmi & Navracsics, 2026) durante los procesos de comprensión y producción orales y escritos.

El desarrollo de la conciencia metalingüística conduce a la profundización de las competencias en lengua extranjera y particularmente de la producción escrita. En la clase de idiomas, una manera de ahondar dicha conciencia es mediante el empleo de los corpus lingüísticos (que suministran datos auténticos de la lengua general o de especialidad) o de los corpus de interlengua como material pedagógico (De Cock & Tyne, 2014). Para ello, es esencial que "el profesor esté familiarizado con las aplicaciones informáticas que posibilitan la búsqueda y extracción de los datos" (Bailini, 2023, p. 20).

Gracias a estos recursos se estudia la lengua a través de la lengua con un aprendizaje consciente, crítico, significativo, independiente o colaborativo. Para esto, es posible implementar el enfoque conocido como Data-Driven Learning que se trata de una forma de aprendizaje "que presenta los datos lingüísticos a los aprendices para ayudarlos a encontrar patrones de uso, de forma inductiva" (Cerda-Canales & Venegas, 2023, p. 484). En esta metodología, los estudiantes se sirven a menudo de herramientas tecnológicas; parten de interrogantes sobre ciertos "mecanismos de funcionamiento del sistema morfosintáctico y léxico-semántico de una lengua extranjera" (Colantino, 2023, p. 77) o de otros fenómenos, examinan las informaciones que se encuentran en los corpus desde un ángulo metalingüístico, las explican y llegan a conclusiones específicas (Bailini, 2023; Cerda-Canales & Venegas, 2023).

Para optimizar el grado de conciencia metalingüística con los corpus de interlengua es igualmente provechoso aplicar algunas pruebas, entre ellas las de juicio de gramaticalidad que facilitan "investigar la adquisición de una segunda lengua en términos de conocimiento implícito y explícito que los estudiantes tienen sobre el sistema de reglas de una lengua extranjera" (Yamamoto, 2023[6]). En estas pruebas se proponen una serie de enunciados a partir de los cuales el estudiante realiza una reflexión sobre la lengua extranjera y en la que a menudo indaga sobre los errores, sus causas y posibles correcciones (Maya-Ruiz, 2024).

3. Metodología

Para este estudio, la metodología privilegiada fue la de la lingüística de corpus. Con el objetivo de indagar sobre los puntos de vista de los estudiantes en cuanto a los errores que pensaban cometer con mayor frecuencia en el aprendizaje del francés y contraponer los datos obtenidos con un análisis de sus producciones escritas, se emplearon dos modalidades de recogida de información, a saber, la constitución de un corpus de aprendientes de niveles iniciales de francés y una encuesta de percepciones.

[6] Citado por Maya-Ruiz (2024, p. 88).

3.1. Constitución del corpus y participantes

En la fase inicial de la recogida de información, se solicitó a los estudiantes de pregrado de la carrera de francés, de una universidad pública en Colombia que cursaban las asignaturas Francés 1 y Francés 2, la contribución con algunos de sus escritos, redactados en el marco de estos cursos. Los 46 estudiantes que desearon participar firmaron un consentimiento informado para tal fin.

Los periodos académicos en los que se recopilaron las producciones escritas fueron los semestres 2023-2, 2024-1 y 2024-2. En el segundo semestre de 2023, 14 de los participantes cursaban la asignatura de segundo nivel; en el primer semestre de 2024, 12 de ellos cursaban la asignatura de primer nivel y en el segundo semestre de 2024, 20 de ellos cursaban la asignatura de segundo nivel. La lengua materna de todos los participantes era el español.

Las producciones escritas se obtuvieron mediante la aplicación de dos exámenes parciales y de un examen final en cada una de las materias. Las instrucciones especificadas en estas pruebas fueron las de leer la tarea, seguir las indicaciones, no emplear herramientas de ayuda, impresas o electrónicas, y respetar el límite de tiempo establecido. Los escritos debían ser redactados en un registro de francés general y cotidiano.

Los criterios lingüísticos de evaluación precisaban respetar las indicaciones gramaticales, de vocabulario, de organización, de presentación, de coherencia, de cohesión y de puntuación. Asimismo, se solicitó un número determinado de palabras o frases en todas las pruebas, excepto en una. Para la asignatura de Francés 1, en el primer parcial se requirió escribir 30 frases, en el segundo no se precisó la cantidad y en el examen final se trataba de 120 palabras. En Francés 2 (2023), los requerimientos fueron de 100 palabras en el primer parcial, de 120 en el segundo y de 150 en el examen final. En Francés 2 (2024), el número de palabras fue siempre de 120.

En lo alusivo a las temáticas de las evaluaciones, para la materia de Francés 1 fueron: redactar tres presentaciones personales en un foro (parcial 1), escribir una lista de compras, una invitación y un itinerario para llegar a un punto de encuentro (parcial 2) y redactar un texto turístico, hablando de los hábitos alimenticios y aconsejando a los extranjeros que querían venir al país (examen final).

Para la asignatura de Francés 2 (2023), los temas fueron: responder un correo electrónico, precisando las actividades hechas y las que se tenían previstas, hablando del clima y describiendo emociones (parcial 1), elaborar un texto descriptivo sobre el antes y el ahora en la vida de una persona, precisando el evento que produce el cambio (parcial 2) y escribir un correo electrónico para restablecer contacto con un amigo de la infancia, evocando recuerdos y solicitando información personal (examen final).

Por último, para Francés 2 (2024), las temáticas fueron: contar un sueño extraño (parcial 1), redactar un artículo de periódico sobre una catástrofe o un evento tecnológico que podría producirse en el futuro, presentando testimonios y describiendo consecuencias sociales, medioambientales y económicas (parcial 2) y contar una aventura de viaje positiva o negativa (examen final).

Los exámenes manuscritos fueron transcritos en formato digital, en un documento .txt. Se efectuaron varias relecturas para verificar que las informaciones originales no habían sido modificadas. El corpus de aprendientes es sincrónico y trasversal; se compone de 126 producciones escritas y de 22 791 palabras.

3.2. Encuesta de percepciones y participantes

Con el propósito de conocer los puntos de vista de los estudiantes universitarios de francés sobre sus errores gramaticales, sintácticos y formales, se elaboró un cuestionario con la herramienta Google Forms. A pesar de que este fue ejecutado en 2025 y se evidencia una limitación metodológica producto del desfase cronológico entre su aplicación y la época de recolección de los datos del corpus de interlengua, se estimó enriquecedor realizarlo por el valor pedagógico que podría aportar, ya que, al conocer las creencias de los estudiantes acerca de sus errores, era posible establecer paralelos con los errores del corpus y profundizar el estudio desde una perspectiva metalingüística.

La primera sección de la encuesta contenía preguntas sociodemográficas sobre la edad, el sexo, los conocimientos previos en francés poseídos antes de haber iniciado la formación universitaria de pregrado y el nivel de lengua que consideraban tener al momento de diligenciar la encuesta.

En la segunda sección del cuestionario se solicitó a los estudiantes indicar la frecuencia ("Nunca", "Rara vez", "Algunas veces", "Frecuentemente" y "Muy frecuentemente") con la que cometían errores de manera general.

En la tercera, cuarta y quinta sección se pidió señalar la periodicidad ("Nunca", "Rara vez", "Algunas veces", "Frecuentemente" y "Muy frecuentemente") con la que incurrían en errores gramaticales (asociados con el uso de determinantes, artículos, sustantivos, pronombres, adjetivos, adverbios, verbos, preposiciones, interjecciones y conjunciones), formales (de ortografía y de puntuación) y sintácticos (relacionados con la omisión y la organización de las palabras).

La siguiente sección contenía una pregunta que solicitaba la frecuencia ("Nunca", "Rara vez", "Algunas veces", "Frecuentemente" y "Muy frecuentemente") con la que revisaban sus escritos antes de entregarlos al profesor. Una pregunta abierta suplementaria pedía precisar las razones por las que no revisaban los escritos antes de entregarlos al profesor.

Antes de remitir el formulario final a los estudiantes, un profesor colaboró valorando su contenido y suministrando sugerencias para adecuarlo de mejor modo a las necesidades de la investigación y al público objetivo. El cuestionario también se compartió a algunos estudiantes quienes expresaron su opinión sobre la claridad de las preguntas.

La encuesta fue enviada al correo electrónico de los estudiantes que habían participado en la constitución del corpus de producciones escritas con sus exámenes parciales y finales. No obstante, teniendo en cuenta que esta podía ser cumplimentada voluntariamente, únicamente 23 de los 46 estudiantes que habían participado anteriormente en la construcción del corpus escrito contestaron el cuestionario. En este punto se presenta otra limitación metodológica, provocada por la disminución del número de participantes. Aun así, el valor de la encuesta sigue siendo útil por el componente metalingüístico que aporta.

3.3. Análisis y resultados de los datos del corpus

El corpus bruto fue etiquetado manualmente. Para la anotación, se tomó en cuenta el protocolo de etiquetado de Louvain (Granger et al., 2022) e igualmente 39 etiquetas de errores (Granger et al., 2022). Dado que este manual se fundamenta en la lengua inglesa, fue necesario efectuar una extensión taxonómica de las etiquetas <GNG> y <GADJG> que aluden al género en el sustantivo y en el adjetivo en francés respectivamente y que no existen en inglés, "a language without grammatical gender" (Długosz & Brown-Bousfield, 2025, p. 2). Las letras de estas etiquetas se fijaron por analogía, siguiendo la lógica de creación

expuesta en el protocolo de etiquetado consultado. Similarmente, se introdujo la etiqueta <GINTERJ> para clasificar los posibles errores alusivos a la categoría gramatical de las interjecciones que no se contempla en dicho manual. La tabla 1 presenta la lista de etiquetas que se emplearon.

Tabla 1

Elementos etiquetados y lista de etiquetas fundamentados ampliamente en el manual de etiquetado de Louvain (Granger et al., 2022)

Errores	Etiquetas	Errores	Etiquetas
Puntuación	<Q>	Adverbio	<GADV>
Ortografía	<FS>	Orden del adverbio	<GADVO>
Determinante demostrativo	<GDD>	Número en verbo	<GVN>
Determinante posesivo	<GDO>	Morfología verbal	<GVM>
Determinante indefinido	<GDI>	Formas no conjugadas	<GVNF>
Otros determinantes	<GDT>	Tiempo verbal	<GVT>
Artículos	<GA>	Auxiliares	<GVAUX>
Sustantivo	<GN>	Categoría gramatical	<GWC>
Número en el sustantivo	<GNN>	Preposiciones dependientes	<X*PR>
Género en el sustantivo	<GNG>	Sustantivo con preposición dependiente	<XNPR>
Pronombre demostrativo	<GPD>	Adverbio con preposición dependiente	<XADVPR>
Pronombre personal	<GPP>	Verbo con preposición dependiente	<XVPR>
Pronombre posesivo	<GPO>	Adjetivo con preposición dependiente	<XADJPR>
Pronombre indefinido	<GPI>	Preposición	<LSPR>
Pronombre reflexivo o recíproco	<GPF>	Omisión	<WM>
Pronombre interrogativo o relativo	<GPR>	Orden errado	<WO>
Orden del adjetivo	<GADJO>	Interjección	<GINTERJ>
Número en el adjetivo	<GADJN>	Conjunciones subordinantes	<LCS>
Género en el adjetivo	<GADJG>	Conjunciones coordinantes	<LCC>
Comparativo, superlativo	<GADJCS>		

Ante la carencia de recurso humano y financiero, el corpus solo fue anotado por una investigadora. Por consiguiente, se prescindió de la fase de doble anotación, del acuerdo interanotador y de la resolución de discrepancias. No obstante, con el fin de garantizar la calidad de las anotaciones en el corpus, además de ajustarse al máximo al protocolo de etiquetado de Louvain, se efectuaron varias revisiones y adecuaciones adicionales.

La herramienta lexicométrica Antconc (Anthony, 2026) facilitó igualmente la verificación y el reconocimiento de otras inconsistencias en las anotaciones. En esta fase se constataron categorías que habían sido mal atribuidas, errores en la digitación de las etiquetas, anotaciones ausentes, errores en las correcciones propuestas, entre otros.

La herramienta informática Antconc (Anthony, 2026) permitió también calcular las categorías y subcategorías con errores más repetidos. Para esto, se incorporó el corpus de aprendientes en este programa y a través de la función KWIC (palabra clave en contexto) se

generaron listas de concordancias que indicaban el número de ocurrencias. En la tabla 2, se exponen las frecuencias absolutas y los porcentajes de errores catalogados en las tres categorías generales de análisis. Del mismo modo, se muestra la densidad de errores por cada 1 000 palabras. Se aprecia que son los errores formales los más representativos.

Tabla 2

Frecuencias absolutas y densidad de errores por cada 1 000 palabras

Categoría	Frecuencias absolutas	Porcentaje	Errores por 1 000 palabras
Errores gramaticales	1 640	39,7 %	71,96
Errores formales	2 254	54,6 %	98,90
Errores sintácticos	237	5,7 %	10,40
Total de errores	4 131	100 %	181,26

El número de frecuencias absolutas y de densidad de errores por cada 1 000 palabras de las subcategorías se presenta en la tabla 3.

Tabla 3

Frecuencias absolutas y densidad de errores de las subcategorías gramaticales, formales y sintácticas

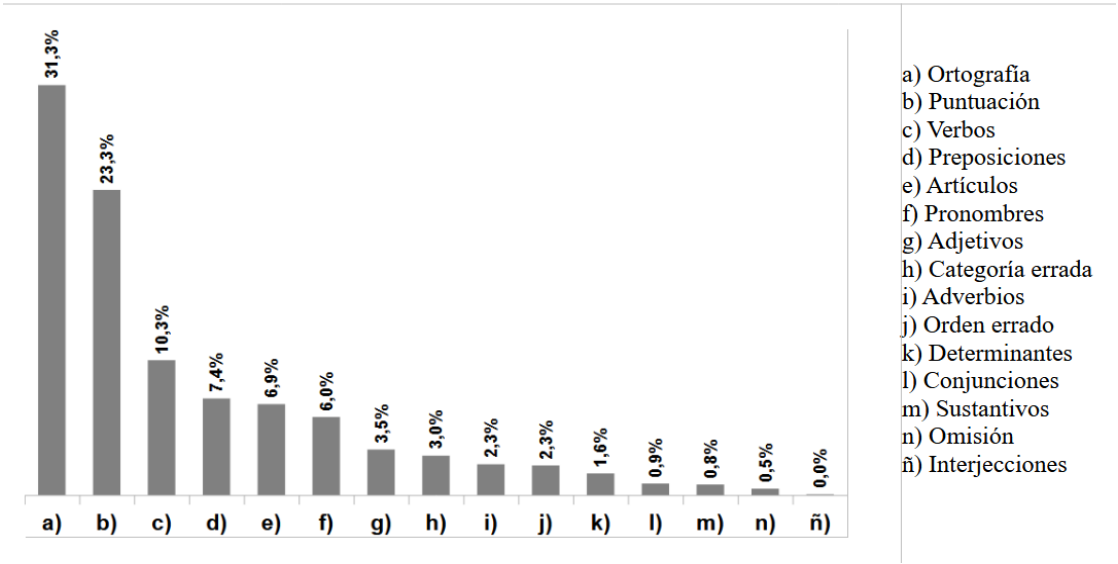
Tipo de error	Frecuencias absolutas	Densidad de errores	Tipo de error	Frecuencias absolutas	Densidad de errores
Determinantes	68	2,98	Adverbios	97	4,26
Artículos	286	12,55	Verbos	425	18,65
Sustantivos	33	1,45	Preposiciones	304	13,34
Pronombres	246	10,79	Interjecciones	2	0,09
Adjetivos	143	6,27	Conjunciones	36	1,58
Puntuación	962	42,21	Ortografía	1 292	56,69
Selección errada	124	5,44	Orden erróneo	93	4,08
Omisión	20	0,88			

Debido al desequilibrio entre el número de producciones escritas en Francés 1 y 2, se optó por no fragmentar el corpus de interlengua en subcorpus, sino más bien analizarlo como un todo compuesto de las producciones escritas de usuarios básicos que comparten rasgos comunes como expresarse de modo sencillo sobre temas en lenguaje cotidiano (Consejo de Europa, 2002) y pese a la realización de tareas heterogéneas.

Los datos contenidos en la tabla 3 muestran que los errores más habituales fueron, en orden descendente, los ortográficos (1 292), los de puntuación (962), los relativos al empleo de los verbos (425), de las preposiciones (304), de los artículos (286), de los pronombres (246), de los adjetivos (143), aquellos relacionados con la elección inadecuada de una categoría gramatical (124), con el uso de los adverbios (97) y con el orden erróneo de las palabras (93). La figura 1 resume los porcentajes de errores del conjunto de subcategorías analizadas.

Figura 1

Porcentajes de errores del conjunto de subcategorías



De los 4 131 errores del corpus, el 31,3 % son ortográficos. Los errores formales ortográficos más constatados aludían al uso del acento agudo "*é*", ya sea por olvido (ej. *ecouté*, decidé*, decouvrais*, hesité*, universite*, amerique*, experience*, itineraire*,* etc.), confusión con el acento grave "*è*" (ej. *derniére*, aprés*,* etc.) o empleo innecesario (ej. *éspere*, més*, éntend*, cartés*, offréra*,* etc.).

Los errores del acento grave con la letra "*è*" aunque observables no fueron tan numerosos; ocurrieron por confusión con el acento agudo (ej. *mèlange*, visitè*, mangè*,* etc.) o por olvido (ej. *collegue, model,* etc.). Otros errores del acento grave, con las vocales como "*ù*" (*aù*, ou*,* etc.) y "*à*" (ej. *apportera*) fueron hallados. Los anteriores no fueron tan predominantes en comparación con un error que se encontró de manera importante entre la preposición "*à*" y el verbo "*a*". Se exponen algunas concordancias: 1) *le cheval a commencé* <*FS corr="à">a</FS> me parler*,* 2) *je vais rester* <*FS corr="à">a</FS> la maison*,* 3) *Tournez* <*FS corr="à">a</FS> droite*.*

El número de errores del acento circunflejo fue reducido, en algunas palabras fueron olvidados (ej. *hotel*, batiment*, chateau*, control*,* etc.) y en otras mal ubicados (ej. *bôite*, goutêr*, disparâtront*, aôut*,* etc.).

Los errores formales ortográficos se produjeron igualmente por la omisión de grafemas. En forma significativa, acaeció con la "*e*" muda al final de las palabras (ej. *j'aim*, a droit*, musé*, planet*, diplom*, ensuit*, mond*, soup*, trist*, control*, pens*, plant*, timid*, construir*, fair*,* etc.). El olvido de la "*e*" muda fue menos repetitivo al interior de las palabras (ej. *normalment*, direcment*, finalment*, seulment*,* etc.). La supresión de la letra "*s*" (ej. *vien*, alor*, mai*, temp*, moin*,* etc.), "*t*" (ej. *clima*, contac*,* etc.) y de otras vocales también fue observada (ej. *beacoup*, hereuse*, amoreux*, acune*, lange*, docter*, geule*, suviens*,* etc.). Asimismo, los errores relativos a los fonemas nasales se apreciaron considerablemente (ej. *commancé*, example*, souvant*, rancontre*, ensamble*, contante*, conférance*,* etc.).

Ahora bien, una cantidad elevada de grafemas erróneos en las palabras resultaron de la influencia interlingüística del español, que es la lengua materna de los participantes, (ej. *functionner*, maravilleux*, mortale*, cientifique*, especial*, estressé*, treinte*, cuatre*, responde*,* etc.) y en un número menos importante, del inglés, que es a menudo una lengua extranjera muy estudiada (ej. *park*, music*, dancer*, japonese*, future*,* etc.). Los fenómenos lingüísticos de la

homonimia y de la paronimia provocaron errores moderadamente (ej. *tente, tante; moi, mois; vue, vu; compagne, campagne,* etc.).

El uso errado de letras mayúsculas y minúsculas se manifestó a menudo, ya sea porque se mezclan dentro de una palabra o frase, porque no se emplean las mayúsculas con los nombres propios (ej. *paris*, italie*,* etc.), porque se emplean las mayúsculas con los meses del año (ej. *Octobre*,* etc.) o porque se confunden con otra categoría gramatical (ej. *colombiens, Colombiens,* etc.). Otros problemas menos habituales corresponden a la mala utilización del guion (ej. *chez-moi*, petit-ami*, aujourd-hui*, vas tu*,* etc.) o a la escritura separada de las letras de una palabra (ej. *bon jour*, a près*,* etc.).

Los errores formales de puntuación ocuparon el 23,3 % de errores del corpus. Los más altamente representativos fueron la ausencia de signos de puntuación, la presencia de espacios antes de una coma o un punto, la utilización inadecuada de un signo de puntuación por otro y el empleo de las comillas inglesas en vez de las francesas. En proporción más baja, se percibió el empleo de signos característicos del idioma español (ej. ¡, ¿) y la presencia de dos signos de puntuación simultáneamente. Algunos ejemplos se presentan enseguida: 1) *C'est pourquoi <Q corr=",">0</Q> je n'ai pas pu aller chez toi ce matin*,* 2) *j'ai eu trois partiels<Q corr=","> ,</Q> c'était difficile !*,* 3) *Ça va <GA corr="les">le</GA>gars<Q corr="?">!</Q>*,* 4) *<Q corr="« je ne veux pas ».">"je ne veux pas".</Q>*,* 5) *<Q corr="0">¿</Q>Comment ça va?*,* 6) *comment <Q corr="vas-tu ?">va-tu?.</Q>*.*

Los errores gramaticales verbales conciernen el 10,3 % del total de errores del corpus. Las subcategorías de errores de tipo verbal más vistas son las relativas a los errores morfológicos (<GVM>; 53,6 %), a la concordancia errada entre el sujeto y el verbo (<GVN>; 18,5 %) y al mal empleo del tiempo (<GVT>; 11,5 %). Los errores que conciernen el uso de los verbos en infinitivo, en gerundio (<GVNF>; 8,4 %) y los auxiliares (<GVAUX>; 7,7 %) fueron menos habituales. Se presentan algunos ejemplos:

1) *<GVM corr="j'étudie">j'étude</GVM> les langues*,* 2) *<GVN corr="j'étais">j'était</GVN> très triste*,* 3) *<GVT corr="j'ai trouvé">je trouvait</GVT> tes photos dans le site web*,* 4) *<GVNF corr="J'aime">J'ai aimer</GVNF>*,* 5) *<GVAUX corr="je me suis réveillée">je me reveillée</GVAUX>*.*

Del total de los errores del corpus, el 7,4 % atañen a las preposiciones. Las preposiciones simples (<LSPR>) poseen un porcentaje superior de error (60,8 %) a diferencia de las que operan conjuntamente (39,1 %) con un verbo (<XVPR>), un sustantivo (<XNPR>), un adverbio (<XADVPR>) o un adjetivo (<XADJPR>). Las preposiciones simples con más alto impacto de error son, de mayor a menor, "*de*", "*à*", "*pour*", "*en*" y "*par*". Las preposiciones que se acompañan de un verbo representan el 77,3 % de dificultad (ej. *parler de, rendre visite à, décider de, commencer à, essayer de, penser à, profiter de, se souvenir de, réfléchir à, sortir de, ressembler à, jouer à,* etc.). Se expone una muestra de frases: 1) *pour promener <LSPR corr="dans">en</LSPR> la ville*,* 2) *<XVPR corr="J'ai essayé de">J'ai essayé</XVPR> dormir*,* 3) *<XVPR corr="sortir de">sortir</XVPR> chez-moi*,* 4) *<XNPR corr="pain au chocolat">pain du chocolat</XNPR>*.*

Del total de 4 131 errores, el 6,9 % y el 6 % alude a los artículos (<GA>) y a los pronombres (<GP>) respectivamente. Los artículos partitivos y los artículos contractos son las construcciones más complejas para los estudiantes.

Los errores de pronombres personales (<GPP>; 37,8%) y de pronombres relativos e interrogativos (<GPR>; 29,6 %) son más usuales respecto a los de pronombres reflexivos y recíprocos (<GPF>; 17 %). Algunas equivocaciones recurrentes categorizadas dentro de los pronombres personales fueron la omisión y la elisión del pronombre personal, el mal uso de los pronombres impersonales y de los pronombres de objeto directo e indirecto. En cuanto a los pronombres relativos e interrogativos, los errores con más prevalencia fueron la formulación de ciertas preguntas cerradas, la confusión entre categorías gramaticales, la

elisión y el uso errado de los pronombres relativos simples y compuestos. En lo que tiene que ver con los pronombres reflexivos y recíprocos, algunos ejemplos de verbos mal empleados fueron: *"promener, marier, voir, coucher, changer, passer, souvenir, rappeler, réveiller, approcher, préparer, tomber, reposer, rencontrer, devenir, rire, trouver,"* etc.

Se presentan algunas concordancias: 1) *<GPP corr="Je m'appelle">J'm'appelle</GPP>*,* 2) *<GPP corr="Je l'ai aidé">Je lui ai aidé</GPP>*,* 3) *<GPP corr="vous">0</GPP> pouvez aller à Bogota*,* 4) *C'est le 26 août, <GPP corr="l'heure">il heure</GPP> 10:00*,* 5) *<GPR corr="qu'est-ce qui se passait ?">qu'est qu'il passait ?</GPR>*,* 6) *ce qui <GPF corr="se passera">passera</GPF> avec eux*.*

Los adjetivos ocupan el 3,5 % de la totalidad de errores del corpus. Los vacíos en el género (<GADJG>; 58,3 %), en el número (GADJN>; 6,29 %) y en el uso de los comparativos y de los superlativos (<GADJCS>; 11,8 %) son los más recurrentes. Se presentan algunos ejemplos: 1) *Il y avait<GADJG corr= "divers">diverses</GADJG>poissons*,* 2) *<GADJCS corr="la meilleure">le meilleur</GADJCS> découverte*,* 3) *la boîte <GADJCS corr="la plus">plus </GADJCS> drôle de la ville*,* 4) *Nous pourrons être <GADJN corr="capables">capable</GADJN>*.*

Los errores relativos a los adverbios (<GADV>; 2,3 %), a los determinantes (<GD>; 1,6 %), a las conjunciones (<LCC> y <LCS>; 0,9 %), a los sustantivos (<GN>; 0,8 %) y a las interjecciones (<GINTERJ>) tuvieron los niveles de incidencia más bajos.

Los errores sintácticos se sitúan con el 5,7 % respecto al porcentaje total de los errores del corpus. El 52,3 % de los errores sintácticos comprenden los errores relativos al uso inadecuado de una categoría gramatical por otra (<GWC>). Se apreció un número elevado de inconsistencias en la elección desacertada entre ciertas categorías. A modo de ilustración, se mencionan las preposiciones y las conjunciones en lugar de verbos (ej. *à/a; et/est; voire/voir*, etc.) o los sustantivos en lugar de conjunciones, verbos o adjetivos (ej. *mai/mais; dos/dois; étude/étudies; sociologie/sociologique*, etc.).

La organización desacertada de las palabras dentro de las frases (<WO>) se hizo visible con un 39,2 % con respecto al conjunto de errores sintácticos, Algunos ejemplos que incluyen el orden errado de adverbios (particularmente de la negación), adjetivos, verbos (sobre todo con el auxiliar, en los tiempos compuestos) y pronombres (más que todo con los pronombres de objeto directo e indirecto) se presentan en una muestra de concordancias: 1) *<WO corr="je n'ai pas vu">j'ai ne vu pas</WO>*,* 2) *<WO corr="Je veux m'y baigner">J'y veux me baigner</WO>*,* 3) *<WO corr="je lui ai dit">j'ai lui dit</WO>*,* 4) *de<WO corr="cuisine française.">française cuisine.</WO>*,* 5) *<WO corr="Le mois prochain">le prochain mois</WO>*,* 6) *<WO corr="Je m'en souviens bien">J'en bien souviens</WO>*.*

3.4. Análisis y resultados de la encuesta de percepciones

De los 23 encuestados, el 60,9 % eran mujeres y el 39,1 % hombres. La mayor parte de ellos tenían predominantemente 19 años (39,1 %) y 22 años de edad (17,4 %). Otras edades moderadamente representativas fueron las de 18, 21, 23 y 24 años (34,8 %) y en menor medida 20 y 30 años (8,7 %).

Respecto a los conocimientos previos en francés, el 73,9 % de los encuestados manifestó no contar con ninguno; el 26,1 % sí poseían algún tipo de competencia en esta lengua.

Por último, en referencia al nivel de lengua que estimaban tener al momento de rellenar la encuesta, el 60,9 % de los estudiantes afirmaron contar con un nivel B1, 30,4 % con un nivel A2 y el 4,3 % con un nivel A1 y B2 respectivamente.

Como se había mencionado previamente, la fecha de realización de las producciones escritas tuvo lugar en los años 2023 y 2024, mientras que la del envío de la encuesta de

percepciones fue efectuada en el año 2025. Por tal razón, se constata un porcentaje más elevado de estudiantes con un nivel B1, ya que avanzaron a un nivel superior de aprendizaje durante este lapso de tiempo.

La frecuencia regular o muy regular en la que las personas interrogadas estiman cometer errores es del 69,6 %; el 30,4 % solo los hacen algunas veces. Los porcentajes de recurrencia en los que se presentan las modalidades de errores gramaticales, formales y sintácticos, se exponen en la tabla 4. Para una mejor lectura, los datos de las categorías "Frecuentemente" y "Muy frecuentemente" fueron fusionados.

Tabla 4

Porcentaje de frecuencia de errores gramaticales, formales y sintácticos

Errores	Nunca	Rara vez	Algunas veces	Frecuente o muy frecuentemente
Determinantes	8,7 %	47,8 %	26,1 %	17,4 %
Artículos	8,7 %	47,8 %	26,1 %	17,4 %
Sustantivos	13 %	47,8 %	26,1 %	13 %
Pronombres	13 %	43,5 %	34,8 %	8,7 %
Adjetivos	17,4 %	34,8 %	21,7 %	26,1 %
Adverbios	13 %	26,1 %	39,1 %	21,7 %
Verbos	4,3 %	13 %	47,8 %	34,8 %
Preposiciones	0 %	13 %	39,1 %	47,8 %
Interjecciones	17,4 %	21,7 %	34,8 %	26,1 %
Conjunciones	13 %	30,4 %	39,1 %	17,4 %
Puntuación	8,7 %	39,1 %	39,1 %	13 %
Ortografía	0 %	4,3 %	21,7 %	73,9 %
Omisión	13 %	26,1 %	34,8 %	26,1 %
Orden erróneo	8,7 %	17,4 %	34,8 %	39,1 %

Para los participantes, los errores más comunes son los ortográficos (73,9 %), los de las preposiciones (47,8 %) y los que tienen que ver con la organización errada de las palabras al interior de las frases (39,1 %). En menor grado de ocurrencia figuran los relacionados con los verbos (47,8 %), los signos de puntuación, las conjunciones, los adverbios (cada uno de ellos con un 39,1 %) y las interjecciones, los pronombres y la omisión de las palabras dentro de las oraciones (cada uno de ellos con un 34,8 %). Aquellos que se observan muy rara vez son los errores relativos al empleo de los determinantes, de los artículos y de los sustantivos (cada uno de ellos con un 47,8 %), seguidos de los pronombres (43,5 %), de los signos de puntuación (39,1 %) y de los adjetivos (34,8 %).

Para terminar, los participantes afirman revisar el texto frecuente o muy frecuentemente, antes de entregarlo al profesor, en un 78,3 %; algunas veces, en un 17,4 % y rara vez, en un 4,3 %. La mayor razón por la que no corrigen los textos es por falta de tiempo y en menor grado por cansancio o falta de motivación.

4. Discusión

En el trabajo realizado por Lopera y Gómez (2021), los últimos resultados revelaron que fueron los errores formales ortográficos y más específicamente, la omisión de los acentos

y las equivocaciones causadas por las divergencias entre los fonemas y los grafemas, los errores más recurrentes cometidos por algunos estudiantes colombianos que aprendían francés en una universidad pública. Al establecer un paralelo con los resultados del presente estudio, se observa una similitud en cuanto a la alta recurrencia de este mismo tipo de errores.

La presencia marcada de los errores ortográficos deriva del carácter opaco de la lengua francesa que se particulariza por la falta de corresponcencia entre la pronunciación y la escritura (De Chambrier et al., 2025; Shen & Del Tufo, 2022). Esto genera confusión y dificultad en los estudiantes, habituados a comunicar en español, una lengua transparente en la que existen divergencias limitadas entre la fonética y la ortografía. "In Spanish, sound-to-print associations are highly consistent, allowing readers to rely on a small grain size for word reading; this is known as a « shallow orthography »" (Marks et al., 2022, p. 2).

Ortografiar correctamente muchos elementos de la lengua francesa implica haber internalizado ciertas reglas morfosintácticas y léxicas como lo son por ejemplo la concordancia de género, de número, las particularidades en las terminaciones de los verbos, las relaciones de las palabras en contexto, las especificidades derivacionales, de homofonía e igualmente determinadas características históricas o etimológicas de este idioma cuyo funcionamiento no posee una explicación lógica en ocasiones (Sénéchal & Gingras, 2017; Sprenger-Charolles, 2023).

En lo que atañe al uso de los determinantes, artículos, verbos y preposiciones simples examinados en el estudio de Viémon (2021) y comparando los porcentajes de aparición de estas categorías con los datos obtenidos en el presente trabajo, se aprecia que los errores en las categorías de los verbos, de las preposiciones y de los artículos comparten rasgos comunes de aparición en el corpus de aprendientes que se constituyó.

El alto predominio de los errores verbales se relaciona igualmente con el trabajo efectuado por Li (2020) en un público de lengua materna china cuyo resultado mostró que los errores más persistentes tenían que ver con la conjugación y la concordancia en un 16 % y con el tiempo, el aspecto y el modo en un 11 %. Asimismo, la gran prevalencia de los errores en los verbos se observa en la investigación realizada por Fong (2022) en un público hispanohablante en el que son una de las categorías más complejas.

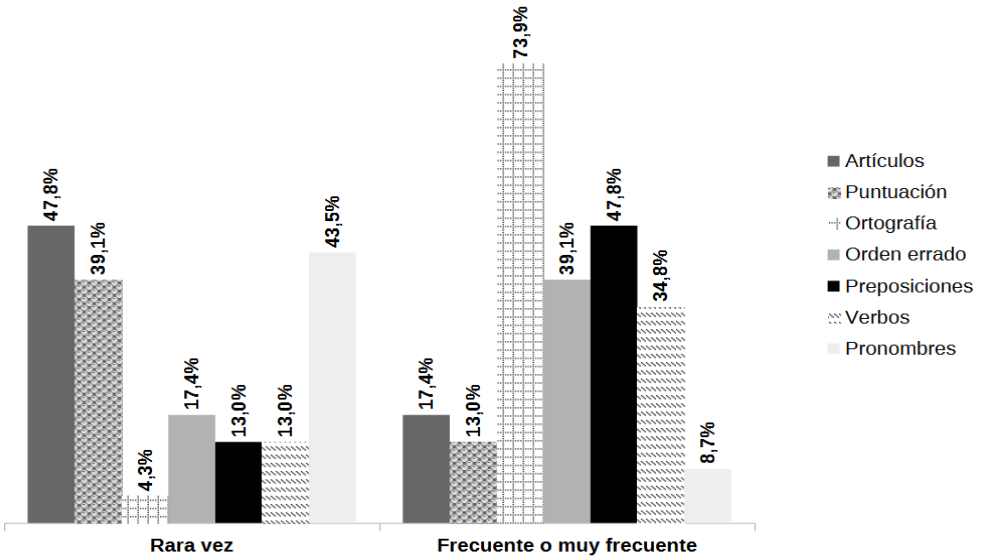
Con respecto a los errores de género que en los trabajos de Viémon (2021) y de Lopera y Gómez (2021) fueron frecuentes y en los que, del mismo modo, se fundamentó la investigación sobre un grupo de discentes germanófonos de Wolf (2025) (y cuyo porcentaje rondaba el 10 %), se pone de manifiesto que en el presente estudio, estos errores se registraron en menor medida.

Ahora bien, con el objetivo de detectar discrepancias o puntos en común entre las creencias de los estudiantes y los errores que cometen en concreto al escribir, se hace necesario triangular los datos obtenidos en la encuesta de percepciones de los estudiantes y en el corpus de interlengua que se diseñó. Para este último, el cálculo de los porcentajes de los errores de las subcategorías fue establecido en función de cada uno de los tres grupos de categorías.

En la figura 2 se resumen los datos de la encuesta de percepciones de las subcategorías en las que los estudiantes piensan cometer más o menos errores y que llaman particularmente la atención.

Figura 2

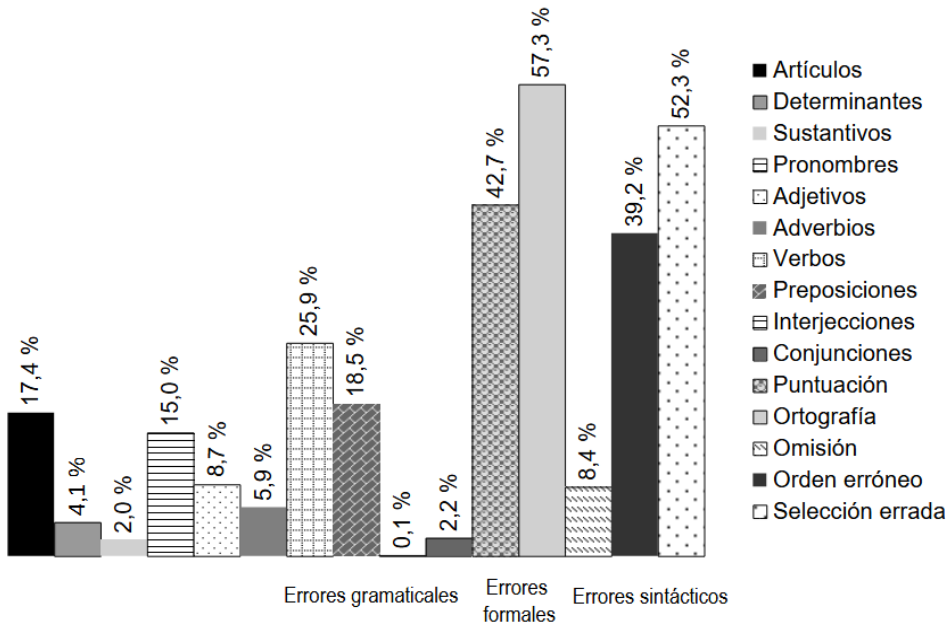
Frecuencia en la que se cometen errores en algunas de las subcategorías según la encuesta de percepciones



En la figura 3 se esquematizan los porcentajes de los errores de las subcategorías, correspondientes al corpus de aprendientes y según la categoría gramatical, formal y sintáctica.

Figura 3

Porcentajes de los errores por subcategoría y según categoría gramatical, formal o sintáctica en el corpus



Se constata que son los porcentajes de los errores formales ortográficos tanto en la encuesta de percepciones (error frecuente o muy frecuente: 73,9 %) como en el corpus

(57,3 %) que se particularizan por poseer la frecuencia más alta de dificultad. Otros porcentajes que concuerdan con una frecuencia alta o muy alta son el de las preposiciones que en la encuesta de percepciones obtuvo un 47,8 % y en el corpus fue la segunda subcategoría gramatical (después de los verbos) que presentó más inconvenientes (18,5 %); y de manera similar, el de la organización errónea de las palabras en las oraciones que en la encuesta de percepciones tuvo un porcentaje de frecuencia o muy alta frecuencia del 39,1 % y en el corpus de aprendientes del 39,2 %.

Por el contrario, la subcategoría de los verbos arrojó resultados un poco opuestos, ya que en el corpus fue el porcentaje con más índice de error (25,9 %), respecto a las otras subcategorías gramaticales, mientras que en la encuesta de percepciones solo fue considerado como un error frecuente o muy frecuente en un 34,8 %. Lo mismo ocurrió con las subcategorías gramaticales de los artículos y de los pronombres que en la encuesta, la gran mayoría de estudiantes, consideraron cometer rara vez (47,8 % para los artículos y 43,5 % para los pronombres), en contraste con los porcentajes del corpus, en los que figuran como la tercera y cuarta subcategorías respectivamente, con los porcentajes de errores más altos después de los verbos y de las preposiciones (17,4 % para los artículos y 15 % para los pronombres). Este fenómeno también se evidenció con la subcategoría formal de la puntuación que en el corpus se presentó con un alto porcentaje de frecuencia (42,7 %) pero en la encuesta obtuvo un porcentaje de frecuencia o de alta frecuencia del 13 %.

Desde una perspectiva metalingüística, se hace patente que los estudiantes reconocen mejor ciertas categorías de errores, mientras que de otras no son tan conscientes de sus falencias. Por tal razón, durante la implementación de las actividades de reflexión, es importante focalizarse aún más en estas categorías que los estudiantes no perciben como problemáticas, pero en las que cometen un número significativo de errores. Lo anterior les permite hacer tareas de reconocimiento, explicar las causas de los errores y proponer soluciones, concienciándolos sobre estos vacíos lingüísticos.

5. Conclusiones

Los errores en los corpus dependen, en cierta medida, de los parámetros establecidos en la realización de las actividades a partir de las cuales se recopila la información. Los datos de las tareas hechas en casa pueden diferir de aquellos que son el resultado de evaluaciones en las que existen restricciones de tiempo y de recursos de aprendizaje. No obstante, aunque las producciones escritas en esta última modalidad reflejen el conocimiento del estudiante, es importante considerar que el proceso de escritura se compone de varias fases como la de revisión y la relectura que son más limitadas en este tipo de ejercicios.

La categoría de errores que se hizo más visible fue la de los errores formales ortográficos y de puntuación. Los primeros se manifiestan, primordialmente, debido a las divergencias significativas que se constatan entre la pronunciación y la escritura de la lengua francesa. Los segundos resultan, de algún modo, de las estrategias cognitivas y lingüísticas que el estudiante implementa durante el proceso de escritura en situación de evaluación, en el que el objetivo es responder a una tarea. En este sentido, da prioridad a ciertos componentes de la lengua como el semántico, el morfosintáctico o el léxico y resta valor a la puntuación por considerarla un elemento secundario. Además, se aprecia que un número importante de estudiantes considera cometer errores de esta índole con poca frecuencia, lo que pone de manifiesto que su conciencia metalingüística no se percata lo suficientemente del rol que estos signos desempeñan en la estructuración de las ideas.

La presencia de determinados errores evidencia el estado de la interlengua de los estudiantes. En este estadio particular, muchas de las estructuras de la lengua aún están en

proceso de internalización o simplemente son desconocidas. Adicional a esto, al escribir un texto durante una evaluación sumativa, el estudiante debe tomar en consideración simultáneamente informaciones de carácter diverso que pueden generar una sobrecarga cognitiva y lingüística, conduciéndolo a escribir erróneamente ciertos elementos. Los errores revelan también los mecanismos que se aplican al tratar de expresarse. Algunos de ellos son la creación de analogías entre la lengua materna, otras lenguas que se manejan y el idioma extranjero que se aprende y la aplicación de generalizaciones o simplificaciones de ciertas estructuras.

Como se mencionó previamente, las semejanzas y las discrepancias entre las respuestas de la encuesta y los datos sobre los errores concretos encontrados en el corpus, muestran que los estudiantes han desarrollado una conciencia lingüística en determinados aspectos pero no en otros.

Con el propósito de brindar una adecuada retroalimentación, es necesario enfocarse en los errores más recurrentes del corpus, identificar sus causas y diseñar tareas de refuerzo. Los corpus de aprendientes son provechosos en el aprendizaje y en la creación de actividades correctivas personalizadas. Al estar constituidos de los propios errores de los estudiantes, reflejan realmente sus vacíos, la manera en la que se sirven de la lengua extranjera y sus necesidades lingüísticas y comunicativas (Tolchinsky, 2014). Algunos ejercicios con extractos del corpus de errores motivan la reflexión y afianzan los conocimientos. Adicionalmente, los estudiantes pueden tratar de encontrar la solución a sus errores o el profesor puede suministrar las correcciones al final del proceso de análisis (Chao & Chacón, 2023).

Las actividades metalingüísticas de mejora deben enfocarse en cada uno de los componentes de la lengua estudiados con el fin de comprender y profundizar en su funcionamiento, matices e interrelaciones. Un ejemplo de ejercicio son los test metalingüísticos en los que a partir de una serie de enunciados, los discentes analizan las reglas que no fueron respetadas y que debieron aplicarse, examinando los motivos que condujeron a los errores desde una perspectiva formal, gramatical, sintáctica o semántica, entre otras, para proponer finalmente posibles soluciones. Es importante resaltar que el desarrollo de la competencia metalingüística conlleva al perfeccionamiento de las habilidades de producion escrita. En efecto, cuanto más profunda sea la reflexión sobre el funcionamiento y la estructuración de la lengua extranjera que se aprende, mayor será la capacidad de escribir diferentes tipos de textos de manera rigurosa (Ghanmi & Navracsics, 2026).

Por último, a pesar de que en su futuro desempeño académico y profesional los estudiantes tienen que contar con las suficientes destrezas lingüísticas, comunicativas y socioculturales en francés, es necesario no valorar los errores como algo completamente negativo. Estos son inherentes al proceso de aprendizaje y es a través de ellos que también se profundizan las competencias.

Declaración de autoría y contribución al artículo

Olga Rocío Serrano-Carranza (primera autora): Conceptualización, curación de datos, análisis formal, investigación, metodología, visualización, redacción del borrador original, redacción, revisión y edición.
Martha Leonor Martínez Acosta (segunda autora): Conceptualización, curación de datos, investigación, metodología.

Referencias bibliográficas

Afifi, Dalia. (2018). L'analyse des erreurs en production écrite chez les apprenants arabophones en classe de FLE. *Anales de Filología Francesa*, (26), 279-293. https://doi.org/10.15446/fyf.v33n1.84182

Agurto, Mónica, Santos, Ofelia, & Chávez, Abdías. (2019). Proceso de la escritura como estrategia para mejorar la producción de textos académicos en estudiantes de bachillerato. *Revista de investigación científic*a, 4(1), 23-28.

Aldama, Nuria, Guerrero, Marta, Montorro, Helena, & Samy, Doaa. (2022). *Anotación de corpus lingüísticos: metodología utilizada en el Instituto de Ingeniería del Conocimiento (IIC)*. Instituto de Ingeniería del Conocimiento.

Anthony, Laurence. (2026). *Antconc* [Computer Software, version 4.4.1]. Waseda University. https://www.laurenceanthony.net/software/antconc/

Bailini, Sonia. (2023). Los corpus de aprendices como recurso didáctico en el aula de español/LE para italófonos. En De Beni, Matteo, Hourani, Dunia, & Sartor, Elisa (eds.), *Comunicación, traducción pedagógica y humanidades digitales en la enseñanza del español como LE/L2/LH* (pp. 19-30). QuiEdit.

Buskivadze, Khatuna, & Baratashvili, Zurab. (2025). Interlingual and intralingual error analysis of Georgian EFL learners. *International journal of multilingual education*, 26(3), 346-360. https://doi.org/10.22333/ijme.2025.10526

Cerda-Canales, Constanza, & Venegas, René. (2023). Uso de corpus para la enseñanza-aprendizaje de la escritura: una revisión sistemática. *Literatura y lingüística*, (48), 479-515. http://dx.doi.org/10.29344/0717621x.48.3448

Chao, Kuok, & Chacón, María. (2023). La retroalimentación correctiva indirecta y la corrección de estudiantes de FLE a nivel universitario. *Revista de filología y lingüística de la Universidad de Costa Rica*, 49(1), 1-32. https://doi.org/10.15517/rfl.v49i1.52284

Colantino, Claudia. (2023). El uso de corpus lingüísticos con fines pedagógicos: el caso de Seah Project. *RLA*, 61(1), 75-92. https://doi.org/10.29393/RLA61-3UCCC10003

Comblain, Annick. (2026). Orthographic depth and spelling development in immersion education: a predictive framework of spelling errors in French. *Languages*, 11, 1-26. https://doi.org/10.3390/languages11060125

Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas. Aprendizaje, enseñanza, evaluación*. Anaya.

Dalle, Laurine. (2024). Identification of written words in dyslexic and non-dyslexic Arabic-speaking learners of L2 French. En Yang, Mimi (eds.), *Multilingualism in its multiple dimensions* (pp. 1-17). IntechOpen.

De Chambrier, Anne-Françoise, Atzemian, Myrto, Fayol, Michel, Zesiger, Pascal, & Martinet, Catherine. (2025). French is not so easy to decode: a pilot study. *Applied Psycholinguistics*, 46(21), 1-20. https://doi.org/10.1017/S0142716425100106

De Cock, Sylvie, & Tyne, Henry. (2014). Corpus d'apprenants et acquisition des langues. *Les cahiers de l'Acedle*, 11(1), 137-168. https://doi.org/10.4000/rdlc.1716

Długosz, Kamil, & Brown-Bousfield, Megan. (2025). The acquisition of grammatical gender in German as an additional language: a comparison of L1 English and L1 Polish learners with and without knowledge of other gendered languages. *Lingua*, 324, 1-18. https://doi.org/10.1016/j.lingua.2025.103989

Ferreira, Anita, & Elejalde, Jessica. (2020). Propuesta de una taxonomía etiológica para etiquetar errores de interlengua en el contexto de un corpus escrito de aprendientes de ELE. *Forma y función*, 33(1), 115-146. https://doi.org/10.15446/fyf.v33n1.84182

Fong, Jesús. (2022). *Representaciones del error en la producción escrita del francés lengua extranjera: una mirada desde el criterio etiológico fundamentado en corpus de estudiantes* [Tesis doctoral, UNED].

García, Juan, & Zuluaga, Juan. (2021). Análisis de la interlengua escrita de estudiantes de nivel intermedio de español como lengua extranjera. *Lengua y habla*, 25, 344-378.

García-Miguel, José. (2022). Lingüística de corpus: de los datos textuales a la teoría lingüística. *Estudios de lingüística del español*, 45, 11-42. https://infoling.org/elies/

Ghanmi, Rabeb, & Navracsics, Judit. (2026). Metalinguistic awareness, writing competence and lexical diversity in bilingual and multilingual adults in Hungary learning English as an additional language. *International journal of multilingualism*, 23(1), 825-849. https://doi.org/10.1080/14790718.2025.2517337

Granger, Sylviane. (2018). Didactique des langues étrangères, linguistique de corpus et traitement automatique des langues. *Les cahiers Fore Llis*. https://cahiersforell.edel.univ-poitiers.fr/index.php?id=595

Granger, Sylviane, Dupont, Maïté, Meunier, Fanny, Naets, Hubert, & Paquot, Magali. (2020). *International corpus of learner English. Version 3*. Presses Universitaires de Louvain.

Granger, Sylviane, Swallow, Helen, & Thewissen, Jennifer. (2022). *The Louvain error tagging manual. Version 2.0*. Centre for English corpus linguistics & UCL Louvain.

Lara, Martha, Egas, Verónica, Rojas, Wilson, & Quishpe, Luis. (2023). Una perspectiva general del aprendizaje basado en errores en la enseñanza de inglés. *Polo del conocimiento*, 8(2), 1816-1842. https://polodelconocimiento.com/ojs/index.php/es/article/view/5276

Li, Jingrao. (2020). *Analyse des erreurs verbales dans un corpus d'apprenants sinophones du français* [Trabajo de maestría, Université Grenoble Alpes].

Lopera, Leidy, & Gómez, Érica. (2021). L'analyse de l'erreur comme stratégie d'intervention de la production écrite en français langue étrangère. *Folios*, (53), 145-170. https://doi.org/10.17227/folios.53-11211

Lozano, Iveth, López, María, & Leon, Vincent. (2025). La fosilización de errores en la interlengua escrita de aprendientes de francés: tratamiento del error. *Ñemitỹrã*, 7(3), 63-74. https://doi.org/10.47133/NEMITYRA20250703b-A5

Macías, Manuel. (2025). Análisis de errores y fracaso pragmático en aprendientes de inglés como LE: evaluación preliminar de resultados en un estudio de caso. *Archivum*, 75(2), 313-352. https://doi.org/10.17811/arc.75.2.2025.313-352

Madrid, Daniel, & Corral, Silvia. (2018). La competencia escrita de alumnos de programas bilingües y no bilingües de educación secundaria. *RMIE*, 23(76), 179-202.

Marks, Rebecca, Sun, Xin, McAlister, Eva, Nickerson, Nia, Hernandez, Isabel, Caruso, Valeria, Satterfield, Teresa, & Kovelman, Ioulia. (2022). Cross-linguistic differences in the associations between morphological awareness and reading in Spanish and English in young simultaneous bilinguals. *International journal of bilingual education and bilingualism*, 25(10), 3907-3923. https://doi.org/10.1080/13670050.2022.2090226

Maya-Ruiz, Selene. (2024). Habilidad metalingüística en inglés como L2 en estudiantes universitarios mexicanos. *Colombian applied linguistics journal*, 26(2), 86-101. https://doi.org/10.14483/22487085.20823

Moreira, María, & Montero, Y. (2025). Competencias lingüísticas en la comunicación escrita de estudiantes de educación básica superior en Ecuador. *Espacios*, 46(1), 99-122. https://doi.org/10.48082/espacios-a25v46n01p08

Quispe-Morales, Rolando. (2022). Desarrollo de la conciencia metalingüística para la comprensión lectora en castellano como segunda lengua. *Educare*, 26(2), 191-209. https://doi.org/10.15359/ree.26-2.11

Real Academia Española. (2025). *Ortografía de la lengua española.* https://www.rae.es/ortograf%C3%ADa/signos-de-puntuaci%C3%B3n

Rodríguez, Victor. (2022). *Escritura académica. El ensayo y su proceso de escritura*. Universidad de la Amazonía.

Sénéchal, Monique, & Gingras, Maxime. (2017). Silex: a database for silent-letter endings in French words. *Behavior research methods*, 49, 1894-1904. https://doi.org/10.3758/s13428-016-0832-z

Shen, Ye, & Del Tufo, Stephanie. (2022). The influence of orthographic depth on multilinguals' neural networks. *Neuropsychologia*, 164, 1-13. https://doi.org/10.1016/j.neuropsychologia.2021.108095

Sprenger-Charolles, Liliane. (2023). Les délires de l'orthographe du français : analyses et propositions. *Pratiques*, 199(200), 1-20. https://doi.org/10.4000/pratiques.14153

Tolchinsky, Liliana. (2014). El uso de corpus lingüísticos como herramienta pedagógica. *Textos de la didáctica de la lengua y la literatura*, (65), 9-17.

Universidad de Antioquia. (2025). *Las cinco fases del proceso de escritura.* https://udearroba.udea.edu.co/externos/mod/hvp/view.php?id=26404#

Viémon, Marc. (2021). Les erreurs fréquentes des étudiants universitaires hispanophones de FLE à l'écrit. *Anales de Filología Francesa*, (29), 769-789. https://doi.org/10.6018/analesff.471371

Wolf, Johanna. (2025). In chunks we trust… the problem of gender assignment in foreign language learning of French. *International journal of bilingualism*, 29(2), 474-490. https://doi.org/10.1177/13670069241302278

Agradecimientos

Se agradece a los estudiantes que desearon cumplimentar la encuesta de percepciones e igualmente a aquellos que brindaron sus producciones escritas con el fin de constituir el corpus de aprendientes.