



ANÁLISIS DE CORPUS TEXTUAL LINGÜÍSTICO EN LOS FOROS ONLINE

ANALYSIS OF LINGUISTIC TEXTUAL CORPUS IN ONLINE FORUMS

Naidelin Silvana Pacheco Chicaiza^{1*}

¹ Estudiante de la Carrera Pedagógica de la Lengua y la Literatura. Universidad Técnica de Cotopaxi. Extensión Pujilí. Ecuador. ORCID: <https://orcid.org/0009-0007-5684-9926>. Correo: naidelin.pacheco7444@utc.edu.ec

Katya Mercedes Grados Fabara²

² Docente Investigador de la Universidad Técnica de Cotopaxi. Extensión Pujilí. Ecuador. ORCID: <https://orcid.org/0000-0002-2708-1256>. Correo: katya.grados8512@utc.edu.ec

* Autor para correspondencia: naidelin.pacheco7444@utc.edu.ec

Resumen

El proceso de estudio se centró en el análisis de un corpus textual-lingüístico generado a partir de un foro en línea con un tema educativo. El objetivo general fue analizar el corpus textual generado en foros online, para mejoramiento del uso del lenguaje en contextos digitales informales. La metodología se basó en un enfoque cualitativo para identificar y caracterizar las estructuras lingüísticas durante la interacción realizada entre estudiante y docente. Además, se basó en un carácter descriptivo, considerando la observación de las actividades desarrolladas mediante el foro, valorando tanto el contenido como el discurso producido. El método empleado en el estudio se basó en el analítico descriptivo, con la finalidad de descomponer el corpus textual en unidades de estudio, como palabras, frases, oraciones y estructuras discursivas, para lo cual se emplearon técnicas como la observación documental y fichas previamente elaboradas para recolectar la información. En conclusión, la investigación evidencia que el lenguaje empleado en los foros educativos realizados en línea se adapta a las dinámicas de inmediatez y funcionalidad comunicativa, constituyéndose como un recurso fundamental para entender las prácticas lingüísticas empleadas en entornos virtuales.

Palabras clave: análisis lingüístico; corpus textual; discurso digital; foros educativos

Abstract

The study focused on analyzing a textual-linguistic corpus generated from an online forum on an educational topic. The overall objective was to analyze the textual corpus generated in online forums to improve language use in informal digital contexts. The methodology employed a qualitative approach to identify and



Esta obra está bajo una licencia Creative Commons de tipo (CC-BY-NC-SA).

Sociedad Ecuatoriana de Investigación Científica. E-mail: revistaalcon@gmail.com



characterize the linguistic structures used in the interaction between students and teachers. It was also descriptive, considering the observation of activities carried out in the forum and evaluating both the content and the discourse produced. The method used in the study was descriptive-analytical, aiming to break down the textual corpus into units of study, such as words, phrases, sentences, and discourse structures. Techniques such as documentary observation and pre-designed data collection forms were used for this purpose. In conclusion, the research shows that the language used in online educational forums adapts to the dynamics of immediacy and communicative functionality, constituting a fundamental resource for understanding the linguistic practices used in virtual environments.

Keywords: *linguistic analysis; textual corpus; digital discourse; educational forums*

Fecha de recibido: 11/10/2025

Fecha de aceptado: 22/12/2025

Fecha de publicado: 15/01/2026

Introducción

Los foros en línea son espacios interactivos de texto escrito donde las personas comparten, construyen experiencias y negocian significados mediante el lenguaje. Estas plataformas se destacan por contener discursos actuales y combinar expresiones coloquiales, usos especializados en educación, recursos basados en argumentación y formas de interactuar de los individuos de forma digital. Asimismo, la confluencia de voces, perspectivas y estilos para comunicarse se han convertido en un aporte valioso para analizar las prácticas lingüísticas actuales, lo que refleja cómo las personas en la actualidad tienden a organizar sus ideas, a sostener argumentos, y desarrollar intercambios discursivos en contextos asincrónicos.

En cuanto a lo manifestado, se da a conocer que el análisis lingüístico en relación con el corpus textual otorga una guía metodológica para estudiar los patrones que surgen a partir de estos espacios virtuales, otorgando la posibilidad de identificar regulaciones léxicas, estructuras intactas, mecanismos de cohesión, estrategias pragmáticas y configuraciones temáticas que permiten conceptualizar la interacción entre participantes, orientando a un presente estudio hacia un enfoque analítico. Además, al proponer un análisis del corpus textual a través de foros educativos, se logra entender las características que respaldan este modelo comunicativo.

El aporte realizado por Vosiljonov (2022), da a conocer que la lingüística de corpus textual es comprendida como una disciplina que encuentra sustento en datos empíricos y se apoya de los avances tecnológicos del siglo XXI, convirtiéndose en un pilar importante para estudiar y desarrollar el lenguaje, específicamente en lo que se refiere al contexto del idioma uzbeko. A través de esta definición, se lo entiende como parte de textos electrónicos, hasta su papel en la creación de diccionarios, la mediación entre frecuencias léxicas y la modelación del lenguaje empleados en la inteligencia artificial, demostrando su importancia a nivel teórico como práctico.



Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).

Sociedad Ecuatoriana de Investigación Científica. E-mail: revistaalcon@gmail.com

Por otra parte, Du et al. (2022), mencionan que el foro educativo online representa una configuración importante y estándar dentro de los recursos, destacando que facilitan establecer interacciones entre estudiantes en tiempo real no obstante dicha participación requiere de una mayor intervención instructiva, así como de la consideración del diseño de plataformas de aprendizaje en línea. Por lo tanto, es importante que los docentes tomen a consideración a la hora de realizar las prácticas educativas, dado que contribuyen al fortalecimiento del diálogo académico.

Además, se detalla que la eficiencia de un corpus textual va a depender del proceso riguroso de selección, digitalización, análisis filológico, normalización y clasificación textual, garantizando la calidad de los datos empleados en enfoques computacionales basados en reglas, aprendizaje autónomo o modelos híbridos. Subrayando que el diseño adecuado de un corpus no solo fortalece la investigación lingüística, sino que además logra impulsar la creación de recursos logísticos que potencian la comunicación, la traducción automática y la proyección internacional de lenguas en procesos de consolidación como el uzbeko.

De acuerdo con Chua (2022), desde los inicios del internet, los debates en línea han proporcionado una ingente cantidad de datos textuales, permitiendo a los investigadores explorar el lenguaje y las prácticas comunicativas a través de este nuevo medio, así como opiniones de los usuarios y la construcción discursiva de los problemas sociales, empleando la metodología lingüística corpus, misma que permite el análisis cuantitativo como el cualitativo, para analizar grandes volúmenes de datos textuales en debates en línea. Ante lo mencionado, Holtz et al. (2022), exponen lo siguiente:

En las últimas décadas, internet ha cambiado la forma en como las personas se comunican, ya sea para comunicación privada o para discusiones sobre temas políticos y sociales. A partir de ello, nace el foro o tablón de mensajes de internet, conocido como un sitio en línea empleado para la discusión. Una de las ventajas a destacar es la cantidad casi ilimitada del material para el análisis, considerando que en algunos foros con miles de usuarios se presentan millones de publicaciones en cientos de miles de hilos (pp. 1-3).

A partir de lo citado, es preciso analizar que los foros en línea generan datos textuales, datos que son considerados naturales, a través de producciones espontáneas de usuarios en contextos de interacción pública o semipública, permitiendo la observación de variaciones en el uso del lenguaje y relacionándolas con la atención comunicativa, audiencia percibida y las convenciones de cada comunidad temática, siendo parte de las características destacables de los foros online.

Investigaciones que han sido abordadas actualmente muestran que los enfoques corpus asistidos combinados con análisis de discurso en la actualidad son altamente efectivos, puesto que, permiten evidenciar y revelar tanto distribuciones cuantitativas como frecuencias léxicas, colocaciones, temas, así como matices cualitativos como tonalidad, evaluación y apelación pragmática encontrados en los textos digitales (Cavasso y Taboada, 2021).

De acuerdo con la investigación desarrollada por Ainhoa (2023), la promesa del aprendizaje basado en ordenadores ha generado impacto en la enseñanza de lenguas y desarrollo de competencias comunicativas, destacando que para el análisis e investigación de foros virtuales se debe considerar diferenciar entre comunicación declarativa (decir), comunicación reactiva (responder) y comunicación interactiva (conversar).



Por el contrario, Wenjing et al. (2020), mencionan que como seres humanos adoptamos consciente e inconscientemente nuestros propios actos comunicativos, sean verbales o no verbales. A través de esta perspectiva de la producción del mensaje, la acomodación comunicativa viene a representar un proceso motivado a nivel social, en el que las personas se adaptan en mayor o menor medida de acuerdo con su propia valoración y su estatus social. Así, la percepción del mensaje, la acomodación comunicativa, suele ser valorada de manera positiva, generando intenciones positivas en futuras interacciones con otras personas, destacando que las personas acuden a foros de apoyo en línea para buscar y brindar apoyo informativo, emocional y de autoestima entre sí, considerando que estos foros se basan en textos, se buscan y ofrecen diversas formas de apoyo a través de este medio.

La investigación de Lawson et al. (2023) evidencian cómo los foros digitalizados se han convertido en un espacio importante para buscar apoyo financiero y emocional cuando las personas se encuentran en situaciones de endeudamiento, específicamente en lo que se refiere a los contextos de crisis económica y precariedad. Es así como, mediante el análisis lingüístico de más de 41 millones de palabras publicadas en estos foros, se ha logrado reflejar que son entornos virtuales altamente impactantes, dado que, aparte de cumplir una función informativa, funcionan como comunidades discursivas donde las personas construyen sus propios significados sobre la percepción que tienen del dinero, la deuda y la vulnerabilidad.

Por su parte, Lobban et al. (2023), muestran que los foros digitales pueden generar efectos diversos en los usuarios, tal es el caso del apoyo emocional, los cuales están estrechamente vinculados con el tipo de interacción lingüística que se produce en ellas. Así, el análisis ha demostrado que la moderación, la estructura del foro, los roles asumidos por los usuarios y la forma en que se desarrolla el intercambio discursivo, influye de manera directa en la calidad comunicativa y en la percepción de apoyo, pertinencia y seguridad, destacando estos espacios, dado que permiten observar patrones recurrentes en el uso del lenguaje, especialmente en lo que se relaciona a expresiones de búsqueda de ayuda, estrategias de acompañamiento textual, manifestaciones de empatía, construcción de identidad y regulación emocional a través de la palabra escrita.

Además, Goźdź (2021), manifiesta que, la revisión del corpus lingüístico evidencia que la lingüística de corpus representa actualmente una herramienta de gran importancia porque permite comprender cómo el lenguaje se estructura, varía y adquiere significados dentro del contexto institucional. El enfoque con el cual parte, analizando el autor, toma en consideración datos que son reales, porque permiten identificar patrones fraseológicos, unidades léxicas que son recurrentes en las plataformas y regularidades discursivas que caracterizan el discurso de las personas, mismos que suelen pasar desapercibidos cuando se emplean métodos tradicionales para analizar el texto. Por ende, la naturaleza comparativa de esta metodología aporta y visibiliza estandarizaciones terminológicas para describir oportunamente los mecanismos lingüísticos argumentativos.

Con referencia a las investigaciones analizadas, se puede reflejar que en la actualidad los foros en línea representan un escenario privilegiado para el análisis lingüístico, considerando que se encuentran basados en corpus. Además, no solo contienen diversidad de datos textuales, sino que también permiten generar una diversidad discursiva que nace de las interacciones entre personas. Así, la literatura concuerda en que estos espacios, aparte de producir un lenguaje natural, espontáneo y contextualizado, contribuyen a la identificación de patrones reales que difícilmente podrían ser observados en entornos que son controlados.



Es importante dar a conocer que, a pesar de la dimisión realizada de los estudios expuestos, el aumento exponencial de los foros en línea, catalogados como espacios de interacción social y discursiva, aún persiste una comprensión limitada sobre los patrones lingüísticos que estructuran la comunicación entre las personas que participan en estos foros. Si bien estos entornos han logrado generar grandes volúmenes de textos variados y espontáneos, la mayoría de las investigaciones han considerado únicamente aspectos funcionales, tales como la participación, moderación y temas tratados, dejando de lado el análisis sistemático del lenguaje empleado desde un enfoque metodológico guiado por corpus textual. Ante ello, la falta de estudios dificulta identificar regulaciones léxicas, programáticas y discursivas que permitan explicar cómo contribuyen las personas, cómo expresan sus posturas, cómo llegan a solicitar apoyo y negocian identidades en un entorno caracterizado por el anonimato y la diversidad sociocultural.

Ante la falta de análisis lingüístico, como bien se ha manifestado, se ha generado una brecha importante entre la comprensión del comportamiento discursivo generado en las comunidades digitales. Ante ello, los foros llegan a representar espacios donde el lenguaje configura dinámicas de interacción que son únicas y cada vez aún más influyentes socialmente. En este sentido, al no contar con herramientas adecuadas y basadas en corpus que permitan el análisis e identificación de patrones reales, resulta complejo lograr la descripción oportuna de la estructura comunicativa basada en intercambios lingüísticos. Por eso, es importante que mediante la presente investigación se realice un estudio usando corpus para identificar los rasgos lingüísticos más importantes de los foros en línea, aportando evidencia empírica que contribuya a la comprensión de cómo se produce, organiza y transforma la comunicación escrita en estos espacios virtuales.

A partir de ello, se plantea la siguiente interrogante: ¿Cuáles son las principales estructuras lingüísticas y características discursivas presentes en los corpus textuales de los foros online, en las plataformas educativas como “Google Classroom” de los estudiantes segundo año de bachillerato de la unidad educativa Gral. de Policía Jorge Poveda, durante el año lectivo 2025-2026?

En este sentido, el objetivo general que guía el desarrollo de la investigación es el siguiente: analizar el corpus textual generado en foros online, para mejoramiento del uso del lenguaje en contextos digitales informales. Para lograrlo se plantean los siguientes objetivos específicos: identificar las categorías lingüísticas más frecuentes en los textos producidos en los foros online; analizar las estructuras sintácticas y léxicas que caracterizan los mensajes escritos en los foros; determinar las funciones discursivas que emergen en las interacciones digitales publicados por los estudiantes.

Materiales y métodos

La presente investigación se realiza considerando un enfoque cualitativo, buscando identificar y caracterizar las estructuras lingüísticas demostradas en la interacción entre estudiantes y docentes, quienes interactúan dentro de los foros en línea con motivos de activar canales de comunicación respecto a los conocimientos y dudas específicas. Es importante mencionar que, este enfoque permite observar los fenómenos como ocurren en su contexto natural sin que exista una intervención por parte del investigador. Así, Valladolid y Chávez (2020), conceptualizan este enfoque como un método que orienta a detallar situaciones y fenómenos tal como estos se presentan, resultando pertinente para analizar el lenguaje en contextos reales de interacción escrita.



El proceso de la metodología para esta investigación es descriptivo, puesto que es importante exponer que, esta aproximación práctica permitirá observar cómo se desarrollará la comunicación escrita en interacciones reales, valorando tanto el contenido lingüístico como el contexto discursivo en el que se producen. Además, la selección del corpus responde a criterios de pertenencia lingüística, riqueza estructural y utilidad comunicativa.

Para desarrollar este estudio, se empleará el método analítico descriptivo, porque permite descomponer el corpus textual en unidades de estudio como palabras, frases, oraciones y estructuras discursivas, mencionando que este método resultará oportuno en la investigación relacionada con datos reales del lenguaje, puesto que permite la observación del comportamiento lingüístico sin idealizarlo. Del mismo modo, se empleó el análisis de contenido, para clasificar los elementos del lenguaje de acuerdo con las categorías gramaticales, semánticas y discursivas.

Por otra parte, se emplearon técnicas de recolección de información que incluyen la observación documental, empleada para obtener textos escritos en los foros en línea y la técnica de análisis textual, porque facilita analizar los fenómenos léxicos, morfosintácticos y semánticos presentados en los mensajes.

La población de este estudio se encuentra conformada por textos producidos en los foros digitalizados que utilizan los usuarios, siendo representada por estudiantes, especialmente los que se encuentran vinculados a ámbitos educativos o de discusión general. Partiendo de ello, la muestra se seleccionará de manera intencional, considerando la relevancia lingüística de los textos, su extensión y la presencia de rasgos discursivos representativos.

Resultados y discusión

En el presente capítulo se dan a conocer los resultados obtenidos mediante la aplicación de los instrumentos de recolección que se han considerado para obtener datos que permitan analizar el corpus textual, tomados de los foros en línea, mencionando que los datos se organizan de manera sistemática lo cual permite un análisis coherente y desarrollado con el fin de entender los fenómenos lingüísticos identificados y las características discursivas presentes en las interacciones escritas por los estudiantes. Cabe mencionar que los resultados se presentarán tanto en gráficas como en tablas para garantizar una mayor comprensión visual e interpretativa.

El foro fue identificado en Reddit, donde profesores y estudiantes opinan con referencia un tema educativo o similares.



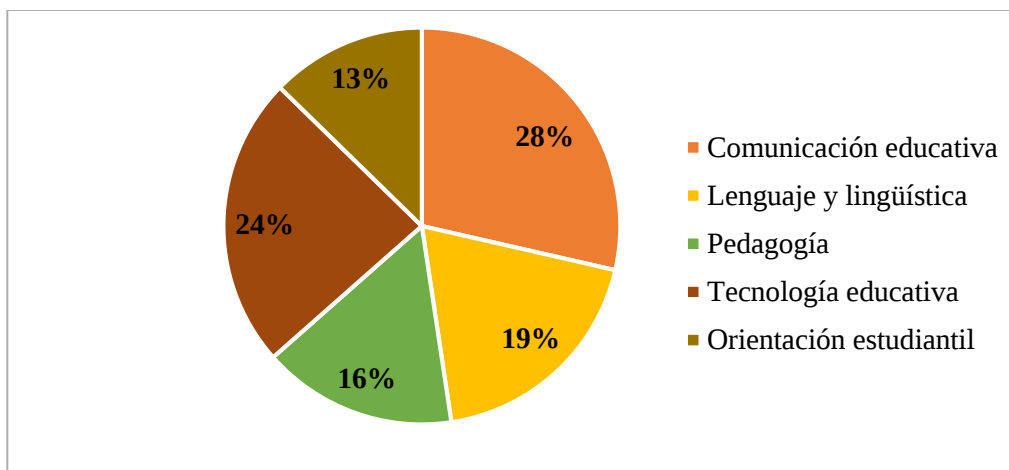


Figura 1. Resultados de la ficha de recolección de corpus textual

En el análisis de la figura 1, se puede demostrar la distribución del corpus mediante una organización sistemática entre disciplinas académicas y profesionales, sumando un total de 63 textos estudiados y analizados. En el ámbito académico, se destaca la comunicación educativa con el 28%, el lenguaje y lingüística con el 19%, y pedagogía con el 16%, develando una clara predominancia de áreas asociadas con el estudio del lenguaje, educación y procesos comunicativos en el contexto de la formación estudiantil. En lo que respecta al ámbito profesional, se destaca la tecnología educativa con el 24% y la orientación estudiantil con el 13%. Ambas disciplinas se asocian con prácticas aplicadas en contextos escolares y digitales.

Los datos estadísticos expuestos contribuyen a anticipar que los fenómenos lingüísticos se encuentran condicionados por la interacción educativa y dinámicas comunicativas propias del contexto pedagógico, tanto en lo teórico como en lo práctico.

Adicionalmente, a través de los datos se puede determinar que la ausencia de textos de comunicación y tecnología educativa expone que es común encontrar estructuras asociadas a la instrucción, explicación y la organización de actividades académicas, en tanto que las disciplinas de lenguaje, lingüística y pedagogía representan un componente técnico conceptual, enriqueciendo la variedad de corpus en términos morfosintácticos y sintácticos.

Tabla 1. Matriz de análisis lingüístico

Nivel léxico (frecuencias, repeticiones, recursos)	Nivel morfosintáctico (tipos de oraciones, conectores, estructuras)	Nivel semántico (temas, campos conceptuales, significados)	Nivel pragmático- discursivo (intenciones, actos de habla, estrategias)	Rasgos del discurso digital (abreviaciones, emojis, marcadores digitales)	Interpretación analítica
Alta frecuencia de frases funcionales (yo, creo, porque,	Predominio de oraciones simples y compuestas cortas. Coordinadas para	Temas asociados a tareas escolares, dudas académicas, instrucciones,	Intenciones principales: solicitar aclaraciones,	Abreviaciones comunes: “xq”, “profe”, “tmb”. Uso moderado de	El discurso evidencia un registro semiformal,



actividad). Repetición de verbos cognitivos (opino, entiendo, necesito). Uso esporádico de tecnicismos académicos.	explicar ideas (y, pero). Subordinadas causales (porque). Escasa complejidad sintáctica.	explicaciones breves y justificaciones. Campos conceptuales: aprendizaje, trabajos, fechas, comprensión.	expresar opinión, justificar respuestas, pedir ayuda. Actos de habla frecuentes: consulta, explicación, agradecimiento.	emojis. Marcadores digitales como “@profe” o referencias a tareas.	mezcla de informalidad digital con funciones académicas. Manifiesta interacción espontánea, lenguaje natural y estrategias básicas de cooperación comunicativa.
Variación léxica limitada, predominio de vocabulario cotidiano.	Uso frecuente de frases verbales con perífrasis modales (tengo que, debo hacer).	Significados orientados a comprensión escolar y manejo de actividades.	Estrategias discursivas centradas en la exposición breve y la economía expresiva.	Presencia de mensajes cortos tipo chat.	La simplicidad del lenguaje refleja la inmediatez propia de foros educativos.
Escaso uso de sinónimos; tendencia a repetir palabras clave de la consigna del foro.	Presencia ocasional de estructuras interrogativas directas cuando solicitan aclaración.	Estructuras interrogativas directas cuando solicitan aclaración. Temas reiterativos: instrucciones, fechas, modos de entrega, contenido curricular.	Uso de mitigadores (creo que, me parece, no estoy seguro) para suavizar postura.	Empleo de signos de exclamación múltiples “!!!”.	El corpus revela patrones de cortesía digital propios de aprendizaje virtual.

La matriz del corpus manifiesta un estilo establecido por el entorno educativo digitalizado, donde la economía expresiva domina cada aspecto interactivo. El vocabulario se mantiene en un registro de la vida diaria, basándose en repeticiones constantes y una mayor preferencia por el uso de palabras funcionales que ayudan a mantener la inmediatez del intercambio. La sintaxis se disminuye a refecciones simplistas o compuestas sin mayor elaboración, sostenidas por conectores básicos que permiten la explicación, justificación o la solicitud de algo sin adornos. Asimismo, los temas giran siempre sobre actividades escolares, dudas recurrentes y explicaciones breves, reforzando la percepción de un discurso centrado en el cumplimiento de tareas.

Tras el comportamiento lingüístico mostrado en la matriz, el foro se configura como un espacio donde prevalece la interacción espontánea, apoyado por la inmediatez estudiantil. Por otro lado, la falta o ausencia de variación léxica y la constante repetición de términos muestran que los participantes tienden a depender en gran medida de un vocabulario de consigna, limitando la diversidad discursiva. No obstante, la presencia de mitigadores interrogativos directos y marcas de cortesía revela una conciencia disminuida sobre la convivencia comunicativa en el entorno digital.

El análisis conjunto de los niveles de lengua, por tanto, permite ver que la puesta en forma del discurso no solo hace eco de decisiones individuales, sino que también responde a las condiciones estructurales en las



que el discurso tiene lugar en el ámbito digital mismo en el que se producen las interacciones. La organización del mensaje, la elección de recursos expresivos y la forma de construir las intervenciones sobre las que se fundamenta la conversación digital son reflejos de la adaptación progresiva a dinámicas comunicativas breves, cortadas, orientadas de forma casi exclusiva a la respuesta inmediata. En este sentido, los mismos elementos destilados del discurso digital, como las abreviaturas, los marcadores gráficos, el formato chat, son soporte de una optimización comunicativa que refuerza la fluidez de la interacción y disminuye la carga cognitiva tanto del emisor como del receptor. Este tipo de discurso propicia una participación enraizada y da continuidad al diálogo, aunque también mediatiza la profundidad argumentativa y la elaboración de conceptos, configurando un modelo de comunicación eficaz que prioriza la cantidad y la rapidez a la complejidad lingüística.

Tabla 2. Guía de observación lingüística

Indicadores	SÍ	NO
Nivel Fonético – Fonológico		
1. Pronuncia adecuadamente los fonemas complejos del español (r, rr, grupos consonánticos).	X	
2. Presenta claridad articulatoria en palabras polisilábicas.	X	
3. Identifica y distingue sonidos similares en producción espontánea.		X
Nivel Morfosintáctico		
4. Construye oraciones simples correctamente (sujeto + verbo + complemento).	X	
5. Utiliza adecuadamente morfemas verbales (tiempos, personas).		X
6. Emplea conectores básicos (“y”, “pero”, “porque”) para unir ideas.	X	
Nivel Semántico		
7. Usa vocabulario pertinente al contexto comunicativo.	X	
8. Selecciona términos con precisión para expresar ideas (sustantivos, verbos y adjetivos adecuados).		X
9. Comprende el significado literal y contextual de palabras clave durante la interacción.	X	
Nivel Pragmático – Discursivo		
10. Mantiene la coherencia temática en intervenciones sucesivas.	X	
11. Infiere información implícita y responde de manera pertinente.		X
12. Respeta turnos conversacionales y adapta su discurso según el interlocutor.	X	
Recursos Lingüísticos Específicos		
13. Emplea marcadores discursivos (“bueno”, “entonces”, “por ejemplo”).	X	
14. Utiliza estrategias de reformulación (“es decir”, “o sea”).		X
15. Manifiesta fenómenos lingüísticos relevante	X	

Se ha seleccionado con relación a los datos e información visualizada en la ficha de observación lingüística, donde se muestra un desempeño favorable en los niveles fonético-fonológico, sintáctico, semántico y pragmático, predominando indicadores positivos. Un alumno presenta una pronunciación adecuada de fonemas complejos y buena claridad articulatoria; no obstante, continúa enfrentando dificultades para diferenciar sonidos similares, reflejando la necesidad de fortalecer la discriminación auditiva. En el nivel morfo-sintáctico, se logran construir oraciones simples de manera correcta, empleando conectores muy básicos, a pesar de que aún persisten limitaciones en el uso apropiado de morfemas verbales. Específicamente



en lo que se refiere a los tiempos y concordancia, afectando la precisión gramatical en producciones más complejas.

En lo que corresponde al nivel semántico, se identifica un adecuado dominio del vocabulario y comprensión del significado literal y contextual de las palabras empleadas. Sin embargo, la elección de términos aún no se percibe consolidada, reduciendo la riqueza y especificidad del discurso. Partiendo del componente pragmático-discursivo, los alumnos consiguen mantener la coherencia temática, respetando los turnos conversacionales. Al mismo tiempo, declaran una adaptación al interlocutor, pero presentan inconvenientes al inferir información implícita, lo que limita la profundidad interpretativa.

Discusión de resultados

En la investigación se identificó que el corpus textual analizado mediante foros educativos se caracteriza principalmente por la ausencia en temas de preferencia de vocabulario funcional y cotidiano. La presencia de la variación léxica y el predominio de estructuras sintácticas simples complementadas por conectores básicos y estrategias discursivas enfocadas a la solicitud de ayuda revelan la facilidad para obtener explicación breve y justificación académica, representando un discurso semi-formativo propio de la inmediatez digital. Estos resultados son similares a los de Kim y Römer (2025), quienes, al emplear un enfoque computacional enfocado directamente en modelos de aprendizaje automáticos, identificación de patrones lingüísticos en foros de *Reddit*, logran mostrar que se hace uso excesivo de pronombres en primera persona, verbos cognitivos y desiderativos. De este modo, la repetición de estructuras léxicas refleja en estados emocionales incertidumbre y búsqueda de apoyo social.

Partiendo de los resultados identificados en el desarrollo investigativo, el análisis del cuerpo textual en foros educativos subraya un discurso principalmente caracterizado por la simplicidad sintáctica, la falta de variación léxica y una orientación funcional asociada al cumplimiento de actividades académicas y dudas generadas en el entorno educativo. Dichas interacciones revelan intenciones comunicativas explícitas, partiendo de realizar solicitudes, aclaraciones, justificar datos, organizar tareas, disminuyendo la ambigüedad discursiva y favoreciendo la coherencia temática. En contraste, la investigación llevada a cabo mediante un foro digital de carácter abierto, como el desarrollado por Barbarestani et al. (2025), donde muestran claramente que existen niveles altos de ambigüedad, subjetividad interpretativa y complejidad pragmática que parten de hacer uso excesivo del sarcasmo, ironía y lenguaje de señalamiento, por lo que en sus resultados confirma que en el ámbito educativo se ha hecho regular el uso del lenguaje, limitando la aparición de problemas discursivos conflictivos.

Los resultados muestran que el corpus textual analizado encuentra su principal característica en el discurso funcional, donde predomina la variación léxica y el uso reiterado de vocabulario empleado en el día a día. Desde el plano semántico y pragmático-discursivo, las fechas de entrega y la comprensión de contenidos junto a estrategias de cortesía digital, mitigadores y economía expresiva representan aspectos vinculados al entorno educativo digitalizado, dando prioridad a la cooperación comunitaria. Estos datos interpretativos se relacionan con lo expuesto por Ananthasubramaniam et al. (2023) quienes por medio de su estudio señalan que en comunidades creadas en la web la concordancia de estilo lingüístico está expuesta al contexto social y dinámico conversacional, impulsando a las personas a adaptarse al vocabulario funcional, pronombres, conectores y estructuras sintácticas mucho más simples al estilo de sus interlocutores.



Adicionalmente, en el desarrollo de la investigación, se evidenció que el discurso productivo a partir de foros educativos llevados a cabo en línea, se caracteriza por una diversa economía expresiva, siendo una de las principales ventajas. Por el contrario, las desventajas de ello radican en la limitada variación léxica y el predominio de estructuras sintácticas mucho más simples con el uso reiterativo del vocabulario usado en la vida diaria, como la presencia de expresiones y conectores funcionales y básicos que se encuentran orientados a la consulta y el cumplimiento de actividades académicas. Estos resultados mantienen relación con los identificados por Basal et al. (2024), quienes mediante su desarrollo investigativo muestran que el corpus facilita que los docentes y estudiantes puedan identificar ciertos patrones empleados en el uso lingüístico, favoreciendo el aprendizaje contextualizado y promoviendo la comprensión funcional y adecuada a partir de estructuras y vocabularios generados mediante diversas situaciones comunicativas.

La forma de comunicación en el contexto del discurso académico en foros educativos digitales se determina por una fuerte finalidad comunicativa, en la que se da cabida a los términos más cotidianos y a las fórmulas más repetitivas vinculadas con la necesidad de una resolución rápida de las dudas relacionadas con el contenido académico. La escasa variación léxica y la mayor inclusión de estructuras oracionales simples con algunos conectores básicos apuntan en uno u otro sentido hacia la tendencia de la economía expresiva y de la claridad informativa. La tendencia a la economía expresiva y a la claridad informativa favorece la rápida comprensión de los mensajes intercambiados y, a su vez, la escasa ambigüedad interpretativa, lo que permite que las interacciones se mantengan en torno a objetivos concretos, la demanda de ayuda, las explicaciones breves o la organización de la actividad académica. En este sentido, el espacio virtual educativo favorece un uso semi-formativo en el que se combinan aspectos de informalidad propios del medio, por un lado y, por el otro, funciones que no son requeridas a nivel académico.

Desde el plano pragmático-discursivo, las interacciones comunicativas específicas y cooperativas referidas se sostienen en estrategias de cortesía digital, atenuadores y marcadores discursivos que rigen la convivencia comunicativa. La coherencia temática se preserva a lo largo de los intercambios, reafirmada en la reiteración de términos vinculados a consignas, fechas y contenidos curriculares. A pesar de que esta regularidad discursiva favorece la estabilidad y la funcionalidad del diálogo académico, lo cierto es que ello también limita la riqueza expresiva y la profundidad interpretativa del discurso. En consecuencia, el lenguaje que predomina en esos espacios se adapta a las demandas del contexto educativo digital, priorizando la eficacia comunicativa y la satisfacción de los propósitos académicos en lugar de la complejidad lingüística o de la exploración discursiva más elaborada.

Conclusiones

Con la investigación realizada se pudo constatar que el corpus textual producido en foros de enseñanza virtual se caracteriza por un uso funcional y de carácter cotidiano, una escasa variedad léxica y estructuras sintácticas simples, principalmente orientadas a la claridad de la comunicación y a las necesidades funcionales que cumplen determinadas finalidades del ámbito académico. Estos rasgos contribuyeron a producir un discurso semi-formativo específico de los entornos digitales de enseñanza, donde la inmediatez, la economía expresiva y la cooperación comunicativa resultan claves; siendo mostrados los tipos de interacción del alumnado hacia un docente, evidenciando a su vez el cumplimiento del objetivo general propuesto.



De la misma manera, el análisis de las estructuras lingüísticas y discursivas pone de manifiesto que las interacciones escritas responden a intenciones de carácter comunicativo evidentes como solicitar aclaraciones, organizar actividades, justificar respuestas o mantener la coherencia temática dentro del mismo foro. Desde una aproximación pragmática y semántica, el uso de las estrategias de cortesía digital, mitigadores y de los marcadores discursivos participa en la regulación del intercambio comunicativo, creando un clima de colaboración y disminuyendo la ambigüedad discursiva. Estos son algunos de los resultados que permiten mostrar la capacidad del lenguaje para adaptarse al contexto educativo digital, destacando así el valor del análisis basado en corpus para la comprensión de las prácticas lingüísticas efectivas.

Por último, este trabajo constituye una pertinente aportación al ámbito de la lingüística aplicada y del conocimiento crítico del análisis de las prácticas del texto digital mediante la presentación de una caracterización sistemática del uso del lenguaje en los foros en el marco de la educación. Se destaca la necesidad de articular metodologías de tipo corpus en el conocimiento del uso de la escritura en contextos virtuales. Desde los resultados de la investigación se propone que futuras investigaciones incluyan aumentar el tamaño del corpus, poner en práctica las comparaciones entre niveles de enseñanza o diferentes plataformas, o conocer las interacciones existentes entre cómo se desarrollan las competencias lingüísticas y el ir participando en foros en la red, con el fin de entender mejor la forma en que contribuyen los foros al proceso de enseñanza – aprendizaje.

Referencias

- Ainhua, E. (2023). Retomar el foro virtual como contexto socio discursivo para el desarrollo de la escritura académico universitaria. *Perspectiva Educativa. Formación de Profesores*, 62(2), 140-164. <https://dialnet.unirioja.es/descarga/articulo/9223311.pdf>
- Ananthasubramaniam, A., Chen, H., & Yan, J. (2023). Explorando la correspondencia de estilos lingüísticos en comunidades en línea: el rol del contexto social y la dinámica de la conversación. *Actas del Primer Taller sobre Influencia Social en las Conversaciones*, 1(10). <https://doi.org/10.18653/v1/2023.sicon-1.7>
- Barbarestani, B., Maks, I., & Vossen, P. (2025). Comprensión y análisis del lenguaje inadecuado en el discurso en línea: un estudio comparativo de anotaciones. *Revista científica de innovacion educativa y sociedad actual*, 2(2). <https://doi.org/10.48550/arXiv.2505.16847>
- Basal, A., Akbas, E., & Bal, B. (2024). Corpus Literacy Training for In-Service English Language Teachers. *Profile: Issues Teach. Prof. Dev*, 26(1). <https://doi.org/10.15446/profile.v26n1.106854>
- Cavasso, L., & Taboada, M. (2021). A corpus analysis of online news comments using the Appraisal framework. *Journal of Corpora and Discourse Studies*, 4(1). <https://jcads.cardiffuniversitypress.org/articles/10.18573/jcads.61>
- Chua, S. (2022). Recopilar y analizar un gran corpus de discusiones en línea para explorar las interacciones de los usuarios. *Lingüística de corpus aplicada*, 2(2). <https://doi.org/10.1016/j.acorp.2022.100017>





- Du, Z., Wang, F., Wang, S., & Xiao, X. (2022). Mejorar la participación de los estudiantes en foros de discusión en línea en cursos online masivos y abiertos: el papel de la participación obligatoria. *Sec. Psicología Educativa*, 13. <https://doi.org/10.3389/fpsyg.2022.819640>
- Goźdz, S. (2021). Lingüística de corpus en el discurso jurídico. *Revista Internacional de Semiótica del Derecho*, 34, 1515–1540. <https://doi.org/10.1007/s11196-021-09860-8>
- Holtz, P., Kronberger, N., & Wagner, W. (2022). Análisis de foros de Internet. *Revistas de Psicología de los medios*, 24(2). <https://doi.org/10.1027/1864-1105/a000062>
- Kim, Y., & Römer, U. (2025). Detection and Analysis of Depression-Related Language in an Online Community: Machine Learning, Topic Modeling, and Corpus-Linguistic Approaches. *Corpus-based Studies across Humanities*, 4(53). <https://doi.org/10.1515/csh-2025-0005>
- Lawson, R., Lutzky, U., Kehoe, A., & Gee, M. (2023). Sorry to hear you're going through a difficult time": Investigating online discussions of consumer debt. *Applied Corpus Linguistics*, 3. <https://www.sciencedirect.com/science/article/pii/S2666799123000163>
- Lobban, F., Coole, M., Donaldson, E., Glossop, Z., & Haines, J. (2023). Improving Peer Online Forums (iPOF): protocol for a realist evaluation of peer online mental health forums to inform practice and policy. *BMJ Open*, 13. <https://doi.org/10.1136/bmjopen-2023-075142>
- Valladolid, M., & Chávez, L. (2020). El enfoque cualitativo en la investigación jurídica, proyecto de investigación cualitativa y seminario de tesis. *Vox juris*, 38(2), 69-90. <https://dialnet.unirioja.es/servlet/articulo?codigo=7628480>
- Vosiljonov, A. (2022). Principios teóricos básicos de la lingüística de corpus. *Academicia Globe: Inderscience Research*, 3(2). <https://doi.org/10.17605/OSF.IO/36RWP>
- Wenjing, P., Feng, B., & Shen, C. (2020). Análisis del capital social, el apoyo social y el uso del lenguaje en un foro en línea sobre la depresión: análisis de redes sociales y de contenido. *Revisa de investigación medica en Internet*, 22(6). <https://www.jmir.org/2020/6/e17365/>

