

Variación formal y funcional de los conectores y operadores discursivos en interacciones digitales: el caso de los comentarios de YouTube en español

Formal and functional variation of discourse connectors and operators in digital interactions: The case of YouTube comments in Spanish

Hernán Robledo

Pontificia Universidad Católica de Valparaíso
hernan.robledo@pucv.cl

Mayra Opazo

Pontificia Universidad Católica de Valparaíso
mayra.opazo.m@mail.pucv.cl

Resumen

Este estudio analiza la variación formal y funcional de conectores y operadores discursivos en comentarios escritos de YouTube, desde un enfoque *corpus-driven* que combina análisis manual y automatizado. Se adopta la perspectiva de la macrosintaxis (Fuentes Rodríguez, 2024a), junto con los conceptos de gramaticalización, construccionalización y construcciones semilibres. El *corpus*, compuesto por más de mil comentarios a videos noticiosos de canales chilenos, permitió identificar 1.110 expresiones con función discursiva, clasificadas según su macrofunción (Fuentes Rodríguez, 2018). A partir de un procedimiento mixto, se proponen 46 paradigmas formal-funcionales que agrupan variantes con una misma función discursiva pero diferente estructura. Los resultados destacan la diversidad formal y alta productividad de conectores interactivos y operadores modales y argumentativos, con múltiples construcciones emergentes aún no fijadas en el sistema lingüístico. Se propone una estrategia metodológica replicable basada en paradigmas y análisis contextual. El estudio aporta a la descripción sistemática de la variación discursiva en redes sociales, con proyecciones hacia la lingüística computacional y la enseñanza del español.

Palabras clave: conectores, operadores, gramaticalización, discurso digital, redes sociales.

Abstract

This study analyzes the formal and functional variation of discourse connectors and operators in written YouTube comments, using a corpus-driven approach that combines manual and automated analysis. The theoretical framework adopts the macrosyntactic perspective (Fuentes Rodríguez, 2024a), along with concepts of grammaticalization, constructionalization, and semi-fixed constructions. The corpus, consisting of over one thousand comments on news videos from Chilean channels, allowed the identification of 1,110 discourse-related expressions classified according to their macrofunction (Fuentes Rodríguez, 2018). Based on a mixed-methods procedure, 46 formal-functional paradigms are proposed, grouping variants that share the same discourse function but differ in structure. The results highlight the formal diversity and high productivity of interactive connectors and modal and argumentative operators, with many emerging constructions not yet fixed in the linguistic system. A replicable methodological strategy based on paradigms and contextual analysis is proposed. The study contributes to the systematic description of discourse variation in social media, with implications for computational linguistics and the teaching of Spanish.

Keywords: connectors, operators, grammaticalization, digital discourse, social media.

1. INTRODUCCIÓN

Las redes sociales y plataformas digitales se han consolidado como espacios esenciales para la circulación de ideas, valores y posicionamientos. Plataformas como YouTube y Twitter/X no solo funcionan como medios de difusión de contenido audiovisual, sino también como entornos altamente interactivos donde los usuarios participan mediante comentarios escritos que combinan expresividad, espontaneidad y funciones argumentativas (Fuentes Rodríguez, 2025; Parini *et al.*, 2025). Estos intercambios, a menudo organizados en HILOS, no pertenecen al ámbito de la escritura formal, sino que reflejan rasgos propios de la oralidad y del registro coloquial del español.

En estos contextos, el uso de recursos lingüísticos con funciones de cohesión y regulación del discurso, como los marcadores del discurso, resulta fundamental. En el marco de la macrosintaxis, se distingue entre conectores y operadores discursivos (Fuentes Rodríguez, 2018, 2020a, 2024a) y su estudio en plataformas digitales resulta especialmente relevante por su papel clave en la construcción de relaciones interpersonales, estructura argumentativa y gestión de la subjetividad (López Quero, 2019; Pano Alamán, 2020). Sin embargo, su identificación y clasificación en estos entornos presenta importantes desafíos teóricos y metodológicos. Uno de ellos es su alta variabilidad formal y funcional, que dificulta su tratamiento tanto manual como computacional.

Tradicionalmente, estas unidades se han descrito como productos de procesos de gramaticalización y lexicalización que las fijan en el sistema lingüístico (Onodera, 2011; Narrog & Heine, 2021; Garachana Camarero, 2024). No obstante, muchas de ellas se encuentran aún en etapas intermedias de fijación, manteniendo cierta flexibilidad estructural al tiempo que cumplen funciones discursivas identificables (Degand & Simon-Vandenbergen, 2011; Fuentes Rodríguez, 2024a, 2024b). Esto permite concebir un continuo de fijación, donde coexisten formas plenamente gramaticalizadas con construcciones emergentes (Pons Bordería, 2013; Grande Alija, 2024).

En investigaciones previas hemos analizado esta variación formal en géneros académicos escritos, como resúmenes y memorias de titulación, identificando PARADIGMAS FORMALES, es decir, agrupaciones de variantes que cumplen la misma función discursiva, con vistas a su clasificación automática (Robledo & Aguilar, 2023, aceptado; Robledo & Rodríguez, 2024, en evaluación). Sin embargo, estos estudios se han centrado en registros formales, dejando fuera géneros más dinámicos, como los comentarios digitales, en los que la expresividad, la interacción y la variabilidad son mayores (Placencia & Parini, 2025).

En este estudio se propone un análisis sistemático de la variación formal y funcional de conectores y operadores discursivos en comentarios escritos de YouTube, desde una perspectiva que integra la macrosintaxis, la gramaticalización, la construccionalización y el concepto de CONSTRUCCIONES SEMILIBRES (Narrog & Heine, 2021; Fuentes Rodríguez, 2024a). Se recolectó un *corpus* que es procesado con técnicas mixtas —manuales y computacionales— para identificar construcciones emergentes con función de conector y operador discursivo, describir su grado de fijación y su comportamiento en la interacción digital.

Desde una perspectiva teórica, el estudio contribuye a la comprensión del cambio lingüístico en curso en el español digital; desde una perspectiva metodológica, aprovecha las propiedades del discurso en redes sociales como un laboratorio natural para observar nuevas formas

discursivas en tiempo real (Tannen & Trester, 2013; Androutsopoulos, 2015; Fuentes Rodríguez, 2024b, 2025), combinando análisis manual con técnicas de procesamiento automático del texto. Los resultados pueden contribuir, además, al modelamiento de géneros digitales a partir de los rasgos discursivos de los textos (Nazar, 2024). Finalmente, sus resultados se pueden aplicar a la enseñanza del español, la lexicografía y el desarrollo de tecnologías lingüísticas (Nazar & Renau, 2023; Nazar *et al.*, 2024; Venegas *et al.*, 2024), aportando un marco empírico útil para el análisis de la cohesión y la coherencia en contextos digitales masivos (Herring *et al.*, 2013).

2. FUNDAMENTOS TEÓRICO-CONCEPTUALES Y ESTADO DE LA CUESTIÓN

2.1. Conectores y operadores discursivos en el marco de la macrosintaxis

Desde la perspectiva de la realización contextual que adopta la macrosintaxis del enunciado (Fuentes Rodríguez, 2018, 2020a, 2020b, 2024a), se propone una alternativa a la noción amplia y ambigua de MARCADORES DEL DISCURSO (Schiffrin, 1987; Martín Zorraquino & Portolés, 1999; Portolés, 2001; Loureda & Acín, 2010), que agrupa unidades muy diversas en cuanto a características y comportamiento discursivo. En su lugar, se distingue entre CONECTORES y OPERADORES DISCURSIVOS, dos categorías diferenciadas en cuanto a las instrucciones procedimentales que señalan, su comportamiento sintáctico y función en la estructura informativa (Fuentes Rodríguez, 2003, 2018).

Los conectores son unidades que han perdido su contenido designativo para adquirir la instrucción de relacionar dos segmentos con indicaciones relativas a la estructuración textual, relación argumentativa, reformulación o exemplificación. En español, su distribución es, con frecuencia, parentética y, en muchos casos, son además polifuncionales –pueden ejercer diferentes funciones, dependiendo del contexto (Herring *et al.*, 2013; Vande Casteele & Fuentes Rodríguez, 2019)– y multidimensionales –en muchas ocasiones añaden a su valor principal otros contenidos en un segundo plano (Fuentes Rodríguez, 2024b). Esto exige un análisis contextualizado que considere tanto la estructura formal como la función pragmático-discursiva de cada ocurrencia. Los conectores se agrupan en subcategorías, según el tipo de relación que establecen entre segmentos: por ejemplo, los conectores aditivos, como *además* o *incluso*, suman información; los opositivos, como *sin embargo* o *en cambio*, introducen contrastes; los concesivos, como *aun así* o *de cualquier forma* presentan excepciones; y los consecutivos, como *entonces* o *por ello*, marcan consecuencias. Existen además conectores reformulativos, temporales, enumerativos, exemplificativos, entre otros, que organizan y jerarquizan la información discursiva.

Los operadores discursivos, por su parte, si bien su marco de funcionamiento es el enunciado, no conectan segmentos ni desempeñan ninguna función con respecto a las oraciones gramaticales. Su contenido procedural refiere a la enunciación, la modalidad, al plano argumentativo o a la estructuración informativa del texto. Al igual que los conectores, pueden ser polifuncionales y multidimensionales, de manera que elementos modales focalizan la información y los enunciativos pueden actuar con función argumentativa (Fuentes Rodríguez, 2003, 2018, 2024a). Los operadores discursivos incluyen expresiones que inciden en el plano de la modalidad, la argumentación o la organización de la información. Entre ellos, se encuentran operadores argumentativos, como *también*, *bastante* o *más*, que refuerzan posturas; modales, como *quizá*, *realmente* u *obvio*, que expresan actitudes; enunciativos, como *en serio* o *de verdad*, que comentan el acto de hablar; e informativos, como *precisamente* o *ya sabes*, que resaltan partes del enunciado (Fuentes Rodríguez, 2018). Estas subcategorías permiten una

clasificación sistemática basada en las instrucciones procedimentales que se transmiten en el discurso.

La distinción entre conectores y operadores resulta clave para el análisis de interacciones digitales, donde proliferan expresiones con usos ambiguos o polifuncionales, que requieren una clasificación precisa según su comportamiento contextual. En el uso de la lengua se observan expresiones que, si bien funcionan como conectores u operadores discursivos, no están del todo fijadas en el español, difieren levemente en su forma, mientras cumplen una misma función (o muy similar) en el discurso. Muchas de estas expresiones se configuran como construcciones semilibres, es decir, mantienen una parte fija –generalmente el núcleo procedural– y una parte variable, como un complemento nominal o una cláusula. Por ejemplo, *no obstante, no obstante ello* y *no obstante esta situación* comparten una función concesiva, pese a sus diferencias formales. Sin embargo, la relación entre forma y función no siempre es unívoca: en ciertos casos, una misma forma lingüística es usada para desempeñar funciones distintas según el contexto comunicativo. Este fenómeno refuerza la necesidad de describir la variación desde un enfoque funcional, que considere tanto la estabilidad relativa de algunas expresiones como su comportamiento flexible y contextual en el discurso escrito en plataformas sociales digitales.

2.2. Construcciones semilibres y procesos de cambio en conectores y operadores discursivos

La invariabilidad formal, considerada tradicionalmente un rasgo prototípico de conectores, operadores y marcadores discursivos (Martín Zorraquino & Portolés, 1999; Portolés, 2001; Pons Bordería, 2006; Elvira, 2015; Fuentes Rodríguez, 2018), es una proyección de la clase gramatical de las partículas al ámbito del discurso (Pavón Lucero, 2003). Así, inventarios y diccionarios especializados, que generalmente incluyen elementos prototípicos de estas categorías, tienden a considerar aquellos que exhiben invariabilidad formal y distribución parentética (Martín Zorraquino & Portolés, 1999; Montolío, 2001; Briz *et al.*, 2008; Fuentes Rodríguez, 2018). Sin embargo, si bien conectores y operadores discursivos provienen de unidades gramaticales, han pasado por procesos de cambio lingüístico graduales, como la gramaticalización, lexicalización, construccionalización y otros (Onodera, 2011; Traugott & Trousdale, 2013; Narrog & Heine, 2021; Deng, 2023; Garachana Camarero, 2024), donde han perdido los rasgos de origen para adquirir otros a nivel discursivo. En este sentido, la propia dinámica de creación de estas categorías permite que los estudios de *corpus* del español frecuentemente adviertan la existencia de estructuras en etapas intermedias de formación, no totalmente fijadas, pero que han adquirido un contenido procedural y podrían fijarse completamente en un futuro como conectores u operadores (Montolío, 2013; Pons Bordería, 2013; Robledo & Nazar, 2018, 2023; Robledo, 2021).

Como advierte Fuentes Rodríguez (2018: 18), se trata de unidades que “[...] están en un proceso vivo de creación de nuevas formas, no sólo para marcar la relación (conectores), sino también, y sobre todo, para expresar las dimensiones discursivas, correspondientes a los operadores”. En este contexto, surge el concepto de construcciones semilibres, que designa expresiones parcialmente fijadas, que combinan elementos fijos y variables, y se caracterizan por su

flexibilidad formal, desemantización parcial, capacidad de encapsular información y de proyectar relaciones discursivas como lo hacen las unidades totalmente fijadas (Fuentes Rodríguez, 2020a, 2024b; Fuentes Rodríguez & González-Sanz, 2024). Por ejemplo, las variantes *de cualquier forma/manera/modo* conservan una función de distanciamiento discursivo independientemente de su estructura variable. Algo similar ocurre con *no obstante esto/eso/ello o para que luego diga/digan*.

La variación y evolución de conectores, operadores discursivos y construcciones semilibres es un proceso complejo que involucra distintos mecanismos de cambio lingüístico: gramaticalización, lexicalización, pragmatalización, construccionalización, entre otros. La gramaticalización implica la pérdida de significado léxico o desemantización, la fijación formal y la dependencia sintáctica de expresiones que originalmente tenían un sentido pleno (Onodera, 2011; Narrog & Heine, 2021). La lexicalización, en tanto, es un proceso de conversión de expresiones previamente sintácticas en unidades léxicas fijas, reduciendo su variación semántica y formal (Brinton & Traugott, 2005). Por ejemplo, la expresión *a pesar de* muestra un alto grado de lexicalización en comparación con variantes como *a pesar de lo anterior* o *a pesar de todo ello*, que reflejan etapas intermedias (Montolío, 2013; Garachana Camarero, 2024). Otros autores describen un proceso de pragmatalización, en el que elementos lingüísticos adquieren funciones interactivas, convirtiéndose en herramientas para organizar la comunicación, más que para describir el mundo (Heine, 2013; Degand & Evers-Vermeul, 2015).

Finalmente, la construccionalización es un proceso de cambio lingüístico enmarcado en la Gramática de Construcciones, donde las unidades léxicas y estructuras sintácticas se integran en asociaciones forma-función que se consolidan en esquemas reutilizables, resultando en una nueva construcción que se integra en la red lingüística de los hablantes (Goldberg, 2006). En el caso de los conectores y operadores discursivos, expresiones originalmente composicionales se reinterpretan y estabilizan como herramientas discursivas, mostrando la capacidad de las lenguas para adaptarse a las necesidades comunicativas de sus hablantes (Traugott & Trousdale, 2013).

Los procesos descritos interactúan de manera compleja en la formación y evolución de conectores y operadores discursivos: la gramaticalización y lexicalización explican los cambios en la forma y significado, mientras que la pragmatalización y construccionalización enfatizan la creación de nuevas estructuras funcionales dentro del sistema lingüístico. Estas dinámicas permiten explicar, por ejemplo, cómo la expresión *sin duda* evolucionó de un significado composicional de ‘falta de duda’ a un operador que enfatiza la certeza o refuerza un argumento (Suárez Hernández, 2023) o cómo la construcción ‘*todo lo que es + [elemento]*’ desarrolló funciones enfáticas para resaltar información relevante (Hoffmann Hölcell, 2023).

2.3. El discurso digital y la interacción en redes sociales

2.3.1. Rasgos del discurso digital y su impacto en la estructuración del discurso

El auge de las plataformas y redes sociales digitales ha transformado profundamente la comunicación, permitiendo que las personas participen cotidianamente en interacciones escritas asincrónicas y multimodales (Placencia & Parini, 2025). En espacios como YouTube y Twitter/X, los usuarios construyen identidades discursivas y negocian significados mediante estrategias propias del entorno digital (Fuentes Rodríguez, 2025). Estos escenarios ofrecen un campo fértil para la investigación lingüística, ya que permiten observar no solo el uso, sino

también el proceso de construcción de la lengua mediante el estudio del uso (Fuentes Rodríguez, 2025).

El discurso escrito en estas plataformas presenta múltiples rasgos que lo aproximan a la oralidad, generando un estilo híbrido entre lo escrito y lo conversacional (Pérez-Sabater, 2024). Se caracteriza por una sintaxis simplificada y concatenada, con oraciones breves y fragmentarias, así como la omisión de artículos o sujetos, lo que refleja el pensamiento inmediato (Briz, 2014). También abundan recursos gráficos para intensificar emociones, como alargamientos vocálicos (*Holaaa*), mayúsculas (*¡NO PUEDO CREERLO!*), signos de puntuación repetidos (!!!; ??) o puntos suspensivos que denotan pausas o complicidad (Vandergriff, 2013; Yus, 2014). Estas estrategias compensan la ausencia de prosodia, reforzando el carácter emocional y expresivo del intercambio digital (Yus, 2014).

Otro rasgo relevante es la reducción textual, mediante abreviaturas (*LOL, DM*), acrónimos y tolerancia a errores tipográficos, los que refuerzan la fluidez y el dinamismo del mensaje (Crystal, 2008). Junto a esto, se incorporan neologismos, coloquialismos y expresiones propias del habla cotidiana, generando un discurso flexible y adaptable al contexto social y temático (Herring, 2004; Briz, 2014).

Estas plataformas también favorecen la conexión interpersonal. Se emplean operadores discursivos propios del habla (como *oye* o *bueno*) y recursos humorísticos o irónicos que refuerzan la identidad grupal (Posteguillo, 2003). Asimismo, los usuarios integran elementos visuales, como *emojis* (Vela Delfa & Cantamutto, 2021) y recurren al cambio de código y expresiones multilingües, lo que evidencia el carácter global y multimodal del discurso digital (Androutsopoulos, 2013; Pérez-Sabater & Maguelouk-Moffo, 2019).

Yus (2010, 2014, 2022) denomina estas prácticas TEXTO ESCRITO ORALIZADO o ALTERACIÓN DEL TEXTO, caracterizadas por grafías fonéticas (“imeil” por *email*), abreviaciones (“thx” por *thanks*), elipsis y otras formas que proyectan inmediatez y cercanía. El predominio de la primera y segunda persona (Yus, 2010) y la incorporación de GRAFICONOS (Sampietro, 2016; Herring & Dainas, 2017; Yus, 2022) refuerzan la dimensión interpersonal y la multimodalidad. Así, la investigación ha pasado de enfocarse en las limitaciones tecnológicas a destacar la diversidad sociolingüística del lenguaje digital (Androutsopoulos, 2013).

Estos rasgos hacen del discurso en plataformas sociales un entorno clave para observar la emergencia y fijación de nuevos mecanismos discursivos. Para este estudio, son fundamentales para comprender la forma, función y variación de conectores y operadores, así como sus procesos de gramaticalización y construccionalización.

2.3.2. Los comentarios de YouTube como espacio discursivo interactivo

YouTube se ha consolidado como un entorno interactivo que permite a los usuarios no solo compartir contenido audiovisual, sino también participar activamente a través de los comentarios escritos. Estos forman parte de una dinámica compleja, en la que confluyen un espacio discursivo audiovisual (video), uno visual (“me gusta”) y otro textual, donde se intercambian opiniones, evaluaciones y reflexiones que enriquecen el diálogo en torno al contenido (Parini *et al.*, 2025).

Los comentarios se organizan en hilos, permitiendo una comunicación tanto individual como grupal, dentro de un marco colaborativo y plural (Parini *et al.*, 2025). Su forma es flexible y estratégica, con recursos como el lenguaje coloquial, apelativos y juicios valorativos, que

facilitan la conexión social y la negociación de significados. Se trata de un espacio textual abierto y dinámico: los comentarios pueden editarse, eliminarse o ampliarse en cualquier momento. Además, el anonimato de los participantes amplía la libertad de expresión, aunque también puede favorecer conductas tóxicas, como el discurso de odio o el ciberacoso, especialmente en temas polémicos (Cheng *et al.*, 2017; Giumetti & Kowalski, 2022; Maruenda-Bataller *et al.*, 2024; Ray *et al.*, 2024; Scheinbaum, 2024).

Desde la perspectiva de este estudio, los comentarios de YouTube constituyen un *corpus* privilegiado para analizar conectores y operadores en español. Su organización en hilos, el fuerte componente argumentativo y la variabilidad estilística permiten observar con nitidez fenómenos de variación formal y funcional, así como procesos de gramaticalización y construccionalización en curso (Kozhukhova & Pastukhova, 2019; Pano Alamán, 2020; De Luca, 2021; Padilla Herrada, 2024).

2.4. Estado de la cuestión

2.4.1. Variación formal de conectores, operadores y unidades afines

La variación formal de conectores, operadores y unidades afines ha sido estudiada desde diversas perspectivas. En el ámbito sociolingüístico, se ha descrito la distribución de variantes con funciones como el control de contacto (*¿cierto?*, *¿me entiendes?*) y los recapitulativos (*al final/ al fin de cuentas*) en el español de Chile (San Martín, 2011, 2015), así como la flexibilidad de marcadores de aproximación y atenuación, como *como (que)* (Sandoval & San Martín, 2020; Repede, 2023). Desde una perspectiva diacrónica, se han abordado procesos de gramaticalización en expresiones como *en plan (de)*, *¿no es cierto?* o construcciones como *de manera/modo/forma/suerte de* (Rodríguez-Abruñeras, 2020; Falusi, 2021; Girón Alconchel, 2022), además de variantes como *digo yo, haz de cuenta (que), ya te/le digo, fíjate (que)* o ‘vaya + SN’ (Guillén Escamilla, 2022, 2023, 2024; Espinal *et al.*, 2024).

Enfoques pragmáticos y dialógicos han explorado la variación de atenuadores como *se ve que*, evidenciales (*¿x?, ¿así de x?*) y expresiones interactivas (*oye, mira, verás*) (Albelda Marco, 2016, 2018; Borreguero Zuloaga, 2017; García Negroni & Libenson, 2020; Jansegers & Albelda, 2020). Desde la macrosintaxis y la construccionalización se han analizado construcciones como *de ahí que* o *para que luego digan*, observando su variación interna y potencial de fijación (Fuentes Rodríguez, 2020b; Grande Alija, 2020). Asimismo, investigaciones recientes han aplicado métodos computacionales e identificado patrones como *cuando sea necesario/de ser necesario/si es necesario o de esa forma/de ese modo/de esta manera* (Robledo, 2021; Robledo & Nazar, 2023).

En conjunto, estos estudios evidencian que la variación formal de conectores y operadores es un fenómeno extendido, pero aún escasamente sistematizado. En este estudio proponemos avanzar en esa dirección, integrando el enfoque de la construccionalización con el análisis de *corpus* reales, en especial de interacciones digitales.

2.4.2. Conectores y operadores en redes sociales digitales

Las interacciones escritas en plataformas digitales constituyen un campo fértil para observar el uso y evolución de conectores y operadores del español. En estos espacios, caracterizados por la espontaneidad, la expresividad y la interacción asincrónica, estas unidades discursivas cumplen funciones clave en la organización del discurso, la construcción de relaciones interpersonales y la expresión de emociones o actitudes (De Piero & Narvaja, 2018; Cantamutto

& Vela Delfa, 2020).

Estudios recientes han documentado la emergencia de operadores procedimentales, como *ve* (Zambrano Ojeda, 2023), *dilo y pregunto* (Uceda Leal, 2023) en Twitter, así como la variación funcional de expresiones afirmativas, como *sí, vale, de lujo y me apunto* (Padilla Herrada, 2023, 2024). En WhatsApp, la expresión *tipo* ha sido descrita como un MARCADOR-MEME multimodal en interacciones juveniles (De Luca, 2021). Otros estudios han analizado operadores modales de atenuación en contextos controversiales de Twitter (Aulit, 2021) y el papel de los marcadores metadiscursivos en la expresión de la ironía en esta misma plataforma (Chen, 2022).

Otras investigaciones, como la de López Quero (2018) han estudiado *venga* como marcador apelativo en Twitter, mientras que Pano Alamán (2020) ha identificado los conectores *pero, además, incluso y pues* como mecanismos frecuentes en comentarios de YouTube. Desde la macrosintaxis, Wiśniewska (2021) destaca la función persuasiva de operadores como *muy, también y más* en la comunidad digital #studygram.

Aun cuando estos aportes son significativos, estudios centrados en la variación formal y funcional de estas unidades en comentarios digitales, orientados a dar cuenta de la sistematicidad de esta variación, son aún escasos. Este estudio se sitúa en esa línea poco explorada, proponiendo un análisis sistemático de su variabilidad estructural, su polifuncionalidad y su grado de fijación en el español de plataformas altamente interactivas como YouTube.

3. METODOLOGÍA

En este artículo presentamos un estudio exploratorio-descriptivo cuyo objetivo es identificar y caracterizar las formas y funciones que cumplen los conectores y operadores discursivos empleados en comentarios de YouTube en español. El análisis de los datos combina los enfoques cualitativo y cuantitativo, técnicas manuales y automatizadas y sigue una aproximación *corpus-driven* al análisis del *corpus* (Tognini-Bonelli, 2001). A continuación, describimos el *corpus* de estudio (apartado 3.1) y los procedimientos de análisis de los datos (apartado 3.2).

3.1. Recolección y composición del *corpus*

Se recolectó un *corpus* de comentarios a videos de canales chilenos de noticias en español desde la plataforma YouTube. Los materiales se descargaron del sitio web de la plataforma (www.youtube.com) y se convirtieron a formato Word para su posterior análisis. Específicamente, el *corpus* de estudio está compuesto por 1.058 comentarios a noticias publicadas en canales de YouTube de medios chilenos, sumando un total aproximado de 22.000 palabras. Los criterios de recolección de los datos fueron: a) relevancia y contingencia de los temas de los videos, consistentes en hechos policiales, como crímenes y asaltos, b) fecha de publicación acotada (agosto de 2024), c) alta cantidad de visualizaciones y d) alta cantidad de comentarios emitidos por los usuarios. La composición del *corpus* se muestra en la tabla 1.

Tabla 1. Composición del *corpus* de estudio.

Número, título y enlace de la noticia	Medio YouTube	Fecha publicación	Cantidad comentarios	Cantidad palabras
1. "ESTO ES CODICIA": Julio César Rodríguez critica a la élite por "Caso Audios" – Contigo en la Mañana https://www.youtube.com/watch?v=2cAONPMORbA	Chilevisión	21 agosto 2024	106	2300

2. ACUSADA AGRESIONES: Claudia Cancino fue destituida como concejal de El Bosque – Contigo en la Mañana https://www.youtube.com/watch?v=ifv7bZ_3wys	Chilevisión	14 agosto 2024	157	2005
3. 11 AUTOS QUEMADOS en violento asalto a empresa Brinks en Rancagua – CHV Noticias https://www.youtube.com/watch?v=Rlk3EyKcS0o	CHV Noticias	16 agosto 2024	213	4600
4. Cae Jauría de Bellavista: Violaron a mujer tras drogarla con siete sustancias – REPORTAJES CHV https://www.youtube.com/watch?v=yujXllGVwEU	CHV Noticias	28 agosto 2024	582	13000
Total			1058	21905

3.2. Procedimiento de análisis del *corpus*

El análisis se organiza en seis etapas secuenciales, combinando herramientas manuales y automáticas para identificar y caracterizar la variación formal y funcional de conectores y operadores discursivos. A continuación, se describen estas etapas.

En primer lugar, se realiza una identificación y etiquetado manual de conectores, operadores y expresiones emergentes con funciones discursivas, conforme a la clasificación del diccionario de conectores y operadores discursivos de Fuentes Rodríguez (2018). El proceso es llevado a cabo por dos analistas, que revisan los siguientes aspectos para determinar la pertinencia de una unidad:

- a) desemantización parcial (de transparencia a gramaticalización),
- b) enriquecimiento pragmático,
- c) variación morfológica en el *corpus*,
- d) pérdida de composicionalidad sintáctica,
- e) aparición entre signos de puntuación o integración en la oración,
- f) presencia de marcas de subjetividad o atención del oyente,
- g) fijación sintáctica parcial,
- h) independencia entonativa o aparición parentética,
- i) ausencia de contenido proposicional,
- j) realización de una función macroestructural o procedimental, aunque la forma no esté incluida en el diccionario.

También se identifican unidades polifuncionales, es decir, formas idénticas que cumplen diferentes funciones discursivas, y se etiquetan según la categoría correspondiente. Para evaluar la fiabilidad de la anotación, se calcula el índice Kappa de Cohen (Artstein & Poesio, 2008; Bayerl & Paul, 2011) tanto en la tarea de identificación como en la clasificación funcional de conectores y operadores. El acuerdo se estima a partir de las coincidencias entre los anotadores en una muestra de 1.110 unidades, con base en los criterios previamente consensuados. En los casos de discrepancia, se discuten las diferencias hasta alcanzar consenso.

En segundo lugar, se cuantifican las ocurrencias totales (*tokens*) por categoría y subcategoría con el objeto de identificar las funciones discursivas más recurrentes y las que presentan mayor diversidad de formas (*types*) en el *corpus*. Para ello, se utiliza el índice de Guiraud (*types*

\sqrt{tokens}), en lugar del simple *type-token ratio* (TTR, *types/tokens*), ampliamente utilizado en estudios de diversidad léxica, pero altamente sensible al tamaño del *corpus* (Tweedie & Baayen, 1998). El índice de Guiraud ha demostrado ser más estable y adecuado para comparar diversidad entre conjuntos de tamaños distintos, ya que compensa parcialmente el sesgo introducido por el número de *tokens* (Malvern *et al.*, 2004; Bestgen, 2024). Además, se establece un umbral mínimo de 10 *tokens* por subcategoría, con el fin de garantizar una mejor representatividad estadística de las muestras y evitar distorsiones producidas por unidades con frecuencias muy bajas. Aunque no existe un consenso universal sobre este umbral, en distintos enfoques del análisis de *corpus* se observan prácticas similares (Rayson, 2008; Baroni *et al.*, 2009; Scott, 2010; Pomikálek, 2011; Baker, 2023). Estos dos enfoques combinados permiten una estimación más fiable en la identificación de las funciones discursivas con mayor diversidad formal.

En tercer lugar, se proponen paradigmas formal-funcionales, definidos como agrupaciones de variantes que comparten una misma función discursiva y presentan una estructura parcialmente común. Estos paradigmas incluyen tanto construcciones semilibres (como *qué bueno*, *Dios quiera*) como unidades de una sola palabra con variación morfológica (por ejemplo, *dado*, *dada*, *dados*, *dadas*). Para su identificación se aplican tres procedimientos complementarios:

- a) Manual, mediante agrupamiento por parte de los analistas.
- b) Automático estructural, a partir de los paradigmas obtenidos manualmente. Se implementan expresiones regulares en Python, diseñadas para detectar patrones formales. Estos paradigmas se validan manualmente para garantizar su pertinencia.
- c) Automático funcional (sinonimia), que considera variantes distintas, pero con una misma función discursiva. Al igual que los paradigmas automáticos estructurales, se basan en los paradigmas manuales e incluyen expresiones regulares implementadas en Python. Se validan manualmente para garantizar su pertinencia.

En cuarto lugar, se identifican los paradigmas más productivos, considerando tanto la cantidad de variantes distintas (*types*) como su frecuencia total en el *corpus* (*tokens*), con el objetivo de observar tendencias de fijación o expansión estructural en el uso discursivo. Se elabora, además, un diagrama de dispersión que permite visualizar los paradigmas más productivos.

En quinto lugar, se extraen ejemplos representativos de cada variante en un contexto ampliado, utilizando el procedimiento KWIC20, que delimita un entorno de 20 palabras a la izquierda y 20 a la derecha de la unidad analizada. Los ejemplos se estructuran en tablas, tras eliminar anotaciones y metadatos, para facilitar su interpretación.

Finalmente, se diseñan fichas analíticas para los paradigmas más representativos. Cada ficha incluye: macrofunción, categoría, subcategoría, posición en el discurso, grado de gramaticalización y construccionalización, observaciones de uso y ejemplos representativos seleccionados a partir de la extracción con KWIC20. Estas fichas permiten observar el comportamiento discursivo y estructural de variantes emergentes o parcialmente fijadas.

4. RESULTADOS Y DISCUSIÓN

Se identificó, clasificó y etiquetó un total de 1.110 ocurrencias (*tokens*) de expresiones con función de conector u operador discursivo, incluidas unidades registradas en el diccionario de Fuentes Rodríguez (2018) y otras no incluidas, pero que cumplen funciones discursivas

equivalentes. Esta ampliación del repertorio permite observar no solo la recurrencia de ciertas formas ya fijadas, sino también la productividad de otras en vías de consolidación. En la tabla 2 se muestra una síntesis de los resultados de la identificación y clasificación de estas unidades. En los gráficos 1 y 2, en tanto, se observa que, entre los conectores, los de adición presentan la mayor frecuencia de uso (*tokens*), pero un repertorio reducido de formas distintas (*types*), sugiriendo una alta fijación de las unidades usadas con esta función. En contraste, los interactivos, aunque menos frecuentes se presentan en muchas formas distintas, sugiriendo una menor fijación de estas unidades. Entre los operadores, los modales y los argumentativos muestran la mayor diversidad formal.

Tabla 2. Síntesis de resultados generales del análisis del *corpus*.

Macrof.	Categoría	Subcateg.	Unidades más frecuentes	Types	Tokens	Frec. relativa (%)	TTR (types / tokens)
Conector	Adición	—	y (59), además (10), al mismo tiempo (3), a propósito (2), aparte (2)	11	84	7,57	0,13
	Concesivo	—	igual (4), aún así (1)	2	5	0,45	0,4
	Conclusivo	—	entonces (5), al final (2), total (2), el resultado (1)	4	10	0,9	0,4
	Consecutivo	—	y (22), por eso (7), así que (4), así (1), de ahí que (1)	10	40	3,6	0,25
	Ejemplificativo	—	así (2), ejemplo (1)	2	3	0,27	0,67
	Justificativo	—	a fin de cuentas (1), de hecho (1), después de esto (1), no por algo (1), total (1)	5	5	0,45	1
	Oposición	—	y (22), eso si (7), en cambio (2), antes que (1), por el contrario (1)	6	34	3,06	0,18
	Ordenador de la materia discursiva	Cierre	etc. (6), y punto (2), chao (1), en fin (1), fin (1)	9	15	1,35	0,6
		Continuativo	y (51), bueno (3), a ver (2), bien (1), de esta forma (1)	6	59	5,32	0,1
		De inicio	y (2), y partiendo (1)	2	3	0,27	0,67
		Enumerativo	por otro lado (1)	1	1	0,09	1
		Interactivos	ojo (5), así es (2), bien (2), bueno (2), hermane (2)	33	45	4,05	0,73
	Recapitulativo	—	y (1)	1	1	0,09	1
	Reformulativo	Concreción	y (1)	1	1	0,09	1
		Corrección	perdón (2)	1	2	0,18	0,5
		Explicación	o sea (5)	1	5	0,45	0,2
	Temporal	—	después (16), mientras (7), luego (5), al principio (1), luequito (1)	7	32	2,88	0,22
Operador	Argumentativo	—	más (74), solo (35), nada (33), también (30), siempre (29)	76	491	44,23	0,16
	Enunciativo	—	claramente (2), con respeto (1), de verdad (1), digo (1), en general (1)	10	11	0,99	0,91
	Informativo	—	pura (9), puros (8), puro (4), es que (2), única (2)	13	35	3,15	0,37
	Modal	—	ojalá (22), jajaja (19), mejor (11), quizás (9), seguro (7)	109	228	20,54	0,48
Totales				310	1110		100

Gráfico 1. Resultados frecuencias (*tokens* y *types*) conectores.

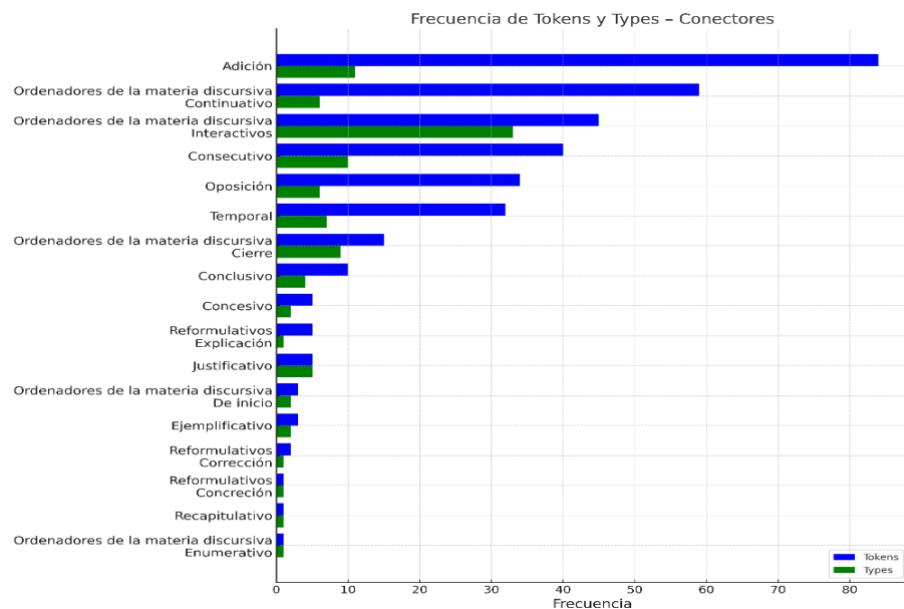
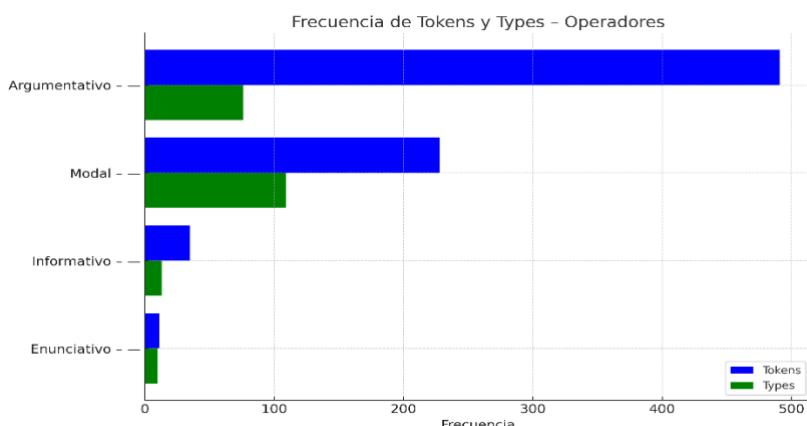


Gráfico 2. Resultados frecuencias (*tokens* y *types*) operadores.



Respecto del acuerdo inter-anotador, en la identificación de conectores y operadores se obtuvo un acuerdo del 85 %, lo que arroja un valor $\kappa \approx 0.60$ (acuerdo moderado, según Landis y Koch, 1977). En la clasificación funcional, el acuerdo observado también fue del 85 %, con un $\kappa \approx 0.79$ (acuerdo sustancial), considerando un 29 % de coincidencia esperada por azar. En ambos casos, las unidades con desacuerdo fueron discutidas hasta consensuar la versión final.

Ahora, como se explica en la subsección 3.2, a partir del cálculo del índice de Guiraud y del establecimiento de un umbral mínimo de 10 *tokens* por subcategoría, se calculó un índice de diversidad formal combinado para identificar las funciones discursivas (macrofunción, categoría y subcategoría) con mayor diversidad de formas distintas en el *corpus*. Los resultados de las categorías que resultaron significativas (≥ 10 *tokens*) se muestran en la tabla 3, mientras que en el gráfico 3 se ofrece una visualización de estas subcategorías.

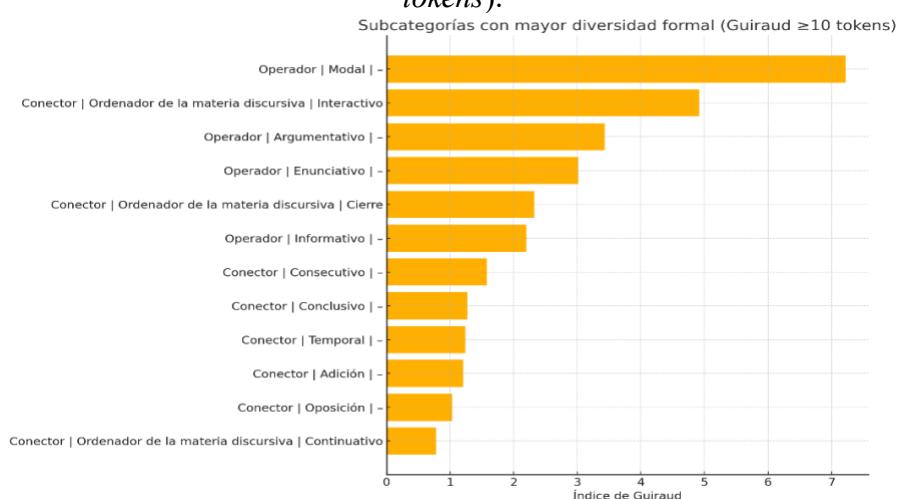
La aplicación de este índice combinado (Guiraud + umbral mínimo ≥ 10 *tokens*) confirma que los operadores modales, ordenadores interactivos y los operadores

argumentativos no solo presentan una alta diversidad formal absoluta, sino también una diversidad relativa significativa, considerando su frecuencia. Asimismo, se observa que ciertas subcategorías menos frecuentes muestran igualmente una notable variación estructural, lo que sugiere procesos activos de fijación y construcción discursiva. Estos hallazgos sugieren que, en la muestra analizada, operadores modales, conectores interactivos y operadores argumentativos constituyen zonas especialmente dinámicas, propensas a la innovación y a la incorporación de formas emergentes, no del todo fijadas en la lengua. Además, refuerzan la pertinencia del enfoque construccional adoptado, que permite sistematizar no solo unidades fijadas, sino también patrones emergentes y variantes formales que participan en la organización discursiva en entornos digitales.

Tabla 3. Índice combinado diversidad formal: Guiraud + umbral mínimo ≥ 10 tokens (solo categorías significativas).

Macrofunción	Categoría	Subcategoría	Tokens	Types	Índice Guiraud	Significativa (≥ 10 tokens)
Operador	Modal	—	228	109	7,22	SÍ
Conector	Ordenador de la materia discursiva	Interactivo	45	33	4,92	SÍ
Operador	Argumentativo	—	491	76	3,43	SÍ
Operador	Enunciativo	—	11	10	3,02	SÍ
Conector	Ordenador de la materia discursiva	Cierre	15	9	2,32	SÍ
Operador	Informativo	—	35	13	2,20	SÍ
Conector	Consecutivo	—	40	10	1,58	SÍ
Conector	Conclusivo	—	10	4	1,26	SÍ
Conector	Temporal	—	32	7	1,24	SÍ
Conector	Adición	—	84	11	1,20	SÍ
Conector	Oposición	—	34	6	1,03	SÍ
Conector	Ordenador de la materia discursiva	Continuativo	59	6	0,78	SÍ

Gráfico 3. Diversidad formal de subcategorías funcionales significativas (basada en índice combinado: Guiraud + umbral ≥ 10 tokens).



En cuanto a unidades polifuncionales, se observó que la forma y desempeña múltiples funciones discursivas en el *corpus*: puede señalar relaciones de adición, continuación y ordenación de inicio, además de funcionar como conjunción copulativa en contextos sintácticos no relevantes para este análisis. Si bien en Fuentes Rodríguez (2018) no se incluye la conjunción copulativa

del español, categorizaciones de marcadores discursivos y unidades afines en la tradición anglosajona sí consideran formas equivalentes como marcas de conexión discursiva. El PDTB-3 (Webber *et al.*, 2019), por ejemplo, incluye *and* como marcador explícito de once relaciones de discurso, siendo la más frecuente la de *expansion.conjunction* en entornos intraoracionales. Esta categoría resulta funcionalmente próxima a las relaciones de conexión de adición descritas en Fuentes Rodríguez (2018), lo que sugiere una convergencia tipológica interesante en el tratamiento de ciertas formas que realizan funciones discursivas en usos específicos.

Con respecto a las variantes formales, construcciones semilibres y paradigmas formal-funcionales, se identificaron 664 ocurrencias (*tokens*) distribuidas en 134 variantes distintas (*types*), organizadas en 46 paradigmas. Estas unidades presentan estructuras parcialmente fijadas, pero funcionalmente recurrentes, y fueron sistematizadas en función de su comportamiento discursivo y su variación estructural. La tabla 4 presenta la propuesta de paradigmas formal-funcionales a partir del *corpus*, cada uno de los cuales incluye: identificador (ID), macrofunción, categoría y subcategoría, origen (manual, automático estructural o automático por sinonimia funcional), variantes, número de variantes y frecuencia total en el *corpus*. Esta sistematización surge de tres procedimientos complementarios: análisis manual de los datos, agrupación automática basada en expresiones regulares que replican la estructura de los paradigmas manuales y agrupación automática por sinonimia funcional, que detecta variantes con comportamientos discursivos similares.

El gráfico 4 sintetiza los resultados de esta propuesta mostrando las macrofunciones, categorías y subcategorías discursivas que presentan mayor VARIACIÓN SISTEMATIZABLE, es decir, aquellas en las que fue posible agrupar distintas variantes bajo un mismo patrón formal-funcional. Los resultados evidencian que los operadores modales y argumentativos, así como los conectores interactivos, destacan por su diversidad y flexibilidad estructural, lo cual coincide con los patrones ya identificados en estudios previos sobre su alta productividad en entornos digitales.

Tabla 4. Propuesta de paradigmas, construcciones semilibres y variantes formal-funcionales de conectores y operadores.

ID	Paradigma propuesto	Macrofunción	Categoría	Subcateg.	Origen	Variantes detectadas	Cantidad variantes	Frec. total (<i>tokens</i>)
1	(así /de ahí) (que)	Conector	Consecutivo	—	Manual	así, así que, de ahí que	3	8
2	(bien)/(bueno)	Conector	Ordenador	Interactivo	Manual	bien, bueno	2	4
3	(mira)/(oye)/(a ver)	Conector	Ordenador	Interactivo	Manual	a ver, mira, oye	3	4
4	(muy/qué) bien	Operador	Argumentativo	—	Manual	bien, muy bien, qué bien	3	3
5	(más) encima	Conector	Adición	—	Manual	encima, más encima	2	7
6	(por) mientras	Conector	Temporal	—	Manual	mientras, por mientras	2	8
7	(qué) lástima	Operador	Modal	—	Manual	lástima, qué lástima	2	3
8	Expresiones con “aún”	Operador	Argumentativo	—	Automático	aún, aún más	2	4
9	Expresiones con “de”	Operador	Modal	—	Automático	de acuerdo, de la nada, de seguro	3	5
10	Expresiones con “dios”	Operador	Modal	—	Automático	dios permita que, dios quiera	2	2
11	Expresiones con “en”	Operador	Enunciativo	—	Automático	en general, en serio	2	2
12	Expresiones con “gracias”	Operador	Modal	—	Automático	gracias, gracias a dios	2	4
13	Expresiones con “hasta”	Operador	Argumentativo	—	Automático	hasta, hasta cuándo	2	15
14	Expresiones con “lo”	Operador	Modal	—	Automático	lo mismo, lo que dios quiera	2	3
15	Expresiones con “más” (sinónimos)	Conector	Adición	—	Automático co sinonimi a	es más, más aún	2	3
16	Expresiones con “más” (sinónimos)	Operador	Argumentativo	—	Automático co sinonimi a	aún más, más o menos, nada más, nadie más	4	4

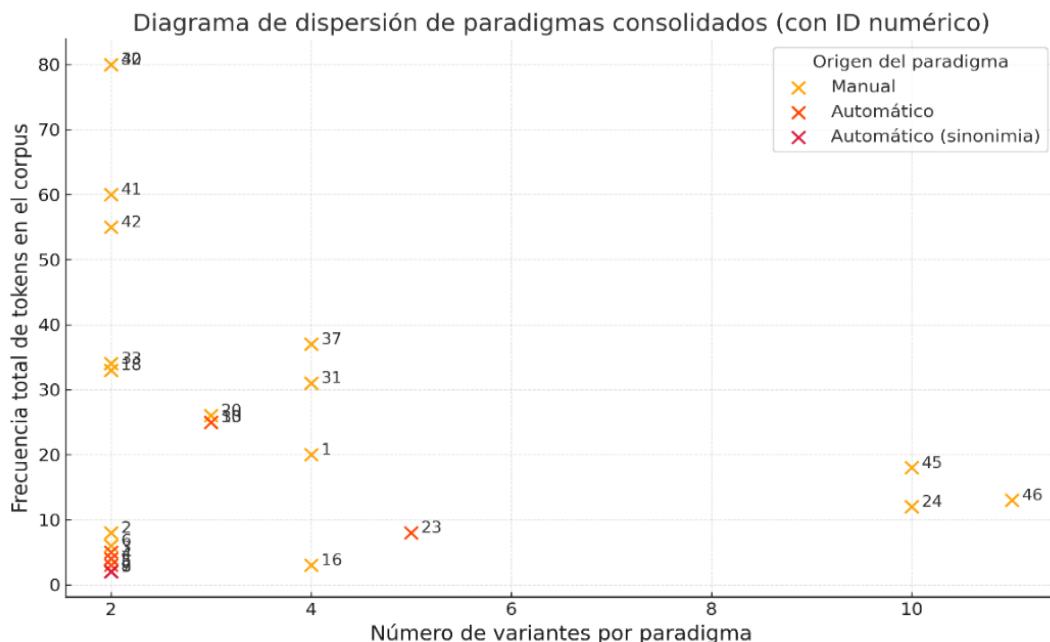
17	Expresiones con “nada”	Operador	Argumentativo	—	Automático	<i>nada, nada más</i>	2	34
18	Expresiones con “nada” (sinónimos)	Operador	Argumentativo	—	Automático sinonimia	<i>nada, nada más</i>	2	34
19	Expresiones con “nadie”	Operador	Argumentativo	—	Automático	<i>nadie, nadie más</i>	2	5
20	Expresiones con “ni”	Operador	Argumentativo	—	Automático	<i>ni, ni menos, ni siquiera</i>	3	25
21	Expresiones con “no”	Operador	Modal	—	Automático	<i>no estoy de acuerdo, no les encanta drogarse?</i>	2	2
22	Expresiones con “por”	Conejor	Consecutivo	—	Automático	<i>por algo, por eso</i>	2	8
23	Expresiones con “por”	Operador	Modal	—	Automático	<i>por ahí, por desgracia, por favor, por fin, por suerte</i>	5	8
24	Expresiones con “qué”	Operador	Modal	—	Automático	<i>qué bueno que, qué estupidez, qué lástima, qué mal, qué manera de, qué mierda de castigo es eso, qué pocos hombres, qué raro que, qué terrible, qué tristeza</i>	10	12
25	Expresiones con “un”	Operador	Argumentativo	—	Automático	<i>un montón, un poco</i>	2	2
26	Expresiones con “y”	Conejor	Ordenador	Cierre	Automático	<i>y chao, y listo, y punto</i>	3	4
27	ah(hh) (chanfle)	Operador	Modal	—	Manual	<i>ah chanfle, ahhh</i>	2	3
28	capaz (que)	Operador	Modal	—	Manual	<i>capaz, capaz que</i>	2	4
29	chu(uu/cha)	Operador	Modal	—	Manual	<i>chucha, chuuu</i>	2	2
30	hart(o/a)	Operador	Argumentativo	—	Manual	<i>harta, harto</i>	2	5
31	much(a/o/a/os/as)	Operador	Argumentativo	—	Manual	<i>muchacha, muchas, mucho, muchos</i>	4	32
32	más ((que)/(y más y más))	Operador	Argumentativo	—	Manual	<i>más, más que, más y más y más</i>	3	79
33	no(más/ más)	Operador	Argumentativo	—	Manual	<i>no más, nomás</i>	2	13
34	nunca (más)	Operador	Argumentativo	—	Manual	<i>nunca, nunca más</i>	2	23
35	obvi(o/amente)	Operador	Modal	—	Manual	<i>obviamente, obvio</i>	2	5
36	ojo (ahí)	Conejor	Ordenador	Interactivo	Manual	<i>ojo, ojo ahí</i>	2	7
37	pur(a/o/os/as)	Operador	Informativo	—	Manual	<i>pura, puras, puro, puros</i>	4	20
38	sobre(todo/ todo)	Operador	Argumentativo	—	Manual	<i>sobre todo, sobretodo</i>	2	4
39	tan(ta)	Operador	Argumentativo	—	Manual	<i>tan, tanta</i>	2	24
40	tod(a/o/os)	Operador	Argumentativo	—	Manual	<i>toda, todo, todos</i>	3	80
41	y (aparte)	Conejor	Adición	—	Manual	<i>y, y aparte</i>	2	61
42	y (después)	Conejor	Ordenador	Continuativo	Manual	<i>y, y después</i>	2	53
43	y (partiendo)	Conejor	Ordenador	De inicio	Manual	<i>y, y partiendo</i>	2	3
44	único(s)	Operador	Informativo	—	Manual	<i>único, únicos</i>	2	2
45	‘qué + ADJETIVO’	Operador	Modal	—	Manual	<i>qué asqueroso, qué bueno, qué finito, qué horribles, qué linda, qué rara, qué raro, qué rasca, qué terrible, qué triste</i>	10	18
46	‘qué + SUSTANTIVO’	Operador	Modal	—	Manual	<i>qué asco, qué bajeza, qué chusmería, qué horror, qué lástima, qué maldad, qué nivel, qué pena, qué picanteño, qué rotería, qué vergüenza</i>	11	13
Totales				M: 27, A: 16, A-S: 3	134 variantes		664 tokens	

Gráfico 4. Funciones discursivas con mayor variación formal sistematizable.



El gráfico 5, por su parte, muestra un diagrama de dispersión que cruza dos dimensiones por paradigma: el número de variantes formales (eje horizontal) y la frecuencia total de ocurrencias en el *corpus* (eje vertical). Cada punto (X) representa un paradigma formal-funcional, identificado por su número de ID. La distribución permite visualizar su productividad y diversidad formal. Los paradigmas en la parte superior derecha combinan alta frecuencia y alta variación, lo que sugiere consolidación y sistematicidad. En la parte superior izquierda predominan formas frecuentes pero poco variables, posiblemente más fijadas. En la zona inferior izquierda se agrupan los menos frecuentes y menos variables, probablemente marginales o emergentes. Finalmente, los situados abajo a la derecha presentan alta diversidad formal pero baja frecuencia, lo que podría reflejar zonas inestables del sistema, vinculadas a procesos de gramaticalización o construccionalización incipiente.

Gráfico 5. Dispersión de paradigmas formal-funcionales consolidados (con ID).



A partir de los 46 paradigmas identificados, se procedió a extraer ejemplos de uso en contexto mediante el procedimiento KWIC20, descrito previamente en la sección metodológica. Este

procedimiento consiste en aislar cada variante formal-funcional en su entorno discursivo inmediato, considerando veinte palabras a la izquierda y veinte a la derecha. La información extraída fue posteriormente analizada para elaborar fichas analíticas detalladas.

En lo que sigue, se presentan cuatro fichas representativas (tablas 5 a 8), correspondientes a algunos de los paradigmas más productivos, frecuentes o polifuncionales a partir del análisis del *corpus*. Estas fichas integran, para cada paradigma, los siguientes elementos: macrofunción, categoría y subcategoría discursiva; posición típica en el discurso; grado de gramaticalización y construccionalización estimado; comentarios de uso basados en el análisis contextual; y ejemplos ilustrativos seleccionados del *corpus*.

La tabla 5 muestra el análisis del paradigma 3 (*mira*, *oye*, *a ver*), clasificado como conector ordenador interactivo. Estas formas, con alto grado de fijación, introducen frecuentemente intervenciones argumentativas o evaluativas dirigidas a otros usuarios, mostrando una función apelativa típica del discurso digital.

Tabla 5. Paradigma 3: (*mira*)/(*oye*)/(*a ver*) como conectores ordenadores interactivos.

Variante	Frec.	Función discursiva	Posición típica	Grado de gramatic.	Grado de construcc.	Observaciones de uso	Ejemplo representativo
a ver	2					Expresiones con función apelativa y de estructuración del turno en discurso digital. Se emplean para iniciar, introducir o redirigir la atención del interlocutor. Uso frecuente en tono irónico o crítico.	... <u>a ver</u> vieja beatriz SON COMO TU DELINCUENTES CHILENOS HERMANITOS TUYOS JAJAJAJA LADRONESSSSSS... [03_107]
mira	2	Conejor ordenador interactivo	Inicio de intervención o de cláusula apelativa	Medio	Bajo a medio		... <u>mira</u> un mapuche violador y delincuente, no eran ángeles?... [04_019]
oye	1						... <u>oye</u> y la gente del motel no hacen nada cuando ven a una tipa ultra dopada con 3 hombres?... [04_027]

La tabla 6, correspondiente al paradigma 9, recoge expresiones con *de* (*de acuerdo*, *de seguro*, *de la nada*) que funcionan como operadores modales con grados variables de gramaticalización. Este paradigma revela cómo ciertas preposiciones con valor semántico pleno han dado paso a construcciones procedimentales.

Tabla 6. Paradigma 9: expresiones con *de* como operadores modales.

Variante	Frec.	Función discursiva	Posición típica	Grado de gramatic.	Grado de construcc.	Observaciones de uso	Ejemplo representativo
de seguro	1					Expresiones modales con distintos matices:	... <u>De seguro</u> eran venezolanos caribeños del tren de Aragua... [04_124]
de acuerdo	6	Operador modal	Inicio o interior de cláusula	Medio	Medio	acuerdo, sorpresa o probabilidad. Algunas pueden	<u>de acuerdo</u> , parece broma, pero es como en el GTA V, en este tipo de asaltos, el asaltante va varios pasos por delante [03_117]

de la nada	3				tener origen en estructuras preposicionales o adverbiales con valor pragmático.	...o mujer. Ahí está la responsabilidad que tiene uno, ser observador, no aceptar weas de otros, no confiar en cualquiera de la nada y con mayor razón andar solo tomar 1 litro de cerveza más pisco. Luego... [04_433]
------------	---	--	--	--	---	---

En la tabla 7, se analiza el paradigma 46 ('qué + sustantivo evaluativo', como *qué vergüenza*, *qué nivel*, *qué rotería*), con usos en contextos de evaluación negativa, sátira o crítica, y una clara función modal expresiva.

Tabla 7. Paradigma 46: construcciones 'qué + SUSTANTIVO' como operadores modales.

Variante	Frec.	Función discursiva	Posición típica	Grado de gramatic.	Grado de construcc.	Observaciones de uso	Ejemplo representativo
qué vergüenza	2						...GRACIAS A DIOS ERA SOLO PARA TV NACIONALFUERA ORDINARIA AHI QUEDASTES TIRADA. LA AUTORIDAD <u>QUE VERGUENZA</u>LA PATA EN LA RAJ... [02_042]
qué bajeza	1						Que bajeza Solo les queda disfrutar lo que votaron Esta es actúa como los delincuentes y... [02_102]
qué lástima	1	Operador modal valorativo	Inicio de oración o exclamación emocional	Medio	Medio		...@sergiofarias6528 PARECIA PELEAS DE CURAOS EN FONDA IGUAL IGUAL @RaulSoto-bf8qd MAÑANA SERA LA ALCANDESA DEL BOSQUE. APUESTAN??? @c.bustamante8205 <u>Qué lástima</u> que quienes son electas autoridades no se compore a la altura Como pueden elegir a estas personas que... [02_037]
qué picanterío	1						...Concejal?, espero alguna respuesta y que tengan una excelente semana . Al mundo jajajajaj <u>Qué picanterío</u> ... Con estas autoridades estamos todos locos... @elybarrera3061 Maravilloso!! Ya era hora que a esta ordinaria la sacaran.... [02_026]

Por último, la tabla 8 presenta el paradigma 41 (*y*, *y aparte*) como conector de adición. Esta construcción, altamente frecuente, refuerza la continuidad temática o el refuerzo argumentativo, integrando una forma básica (*y*) con una expansión (*y aparte*) que evidencia procesos de fijación parcial.

Tabla 8. Paradigma 41: *y (aparte)* como conectores de adición.

Variante	Frec.	Función discursiva	Posición típica	Grado de gramatic.	Grado de construcc.	Observaciones de uso	Ejemplo representativo
y	61	Conejor de adición	Inicio o interior de cláusula	Alto	Alto	Es la forma más polifuncional del <i>corpus</i> y altamente fijada. En función aditiva se presenta como	Esa mujer es narco, en la población todos lo saben, por eso se cree intocable, menos mal que la

					conector mínimo, frecuentemente entre enunciados que comparten orientación temática o argumentativa.	sacaron, <u>y</u> no se olviden de investigar más a fondo, hay sorpresas todos los vecinos saben... [02_076]	
y aparte	2	Conejor de adición	Interior de cláusula	Medio	Medio	Amplifica el valor aditivo e incorpora un matiz valorativo. Su uso es expresivo y refuerza la argumentación evaluativa con énfasis o intensificación. Puede considerarse una construcción semilibre con núcleo fijo y expansión léxica que refuerza la adición. Muestra desemantización parcial y orientación procedural clara.	...puros delincuentes los dejan entrar con armas droga <u>y aparte</u> les cobran para que los dejen vender droga a dentro y los guardias puros extranjeros... [04_185]

Estas fichas analíticas ilustran cómo se concreta la variación formal y funcional en unidades que operan como conectores y operadores en el discurso digital. Además, refuerzan la pertinencia del enfoque construccional adoptado, al evidenciar que junto a formas plenamente fijadas coexisten construcciones emergentes con grados variables de fijación y sistematización.

En conjunto, estos hallazgos amplían las descripciones tradicionales centradas en marcadores discursivos estabilizados y proponen una vía para integrar patrones emergentes dentro de una tipología dinámica y sensible a la variación.

5. CONCLUSIONES

Este estudio exploratorio-descriptivo ha analizado la variación formal y funcional de conectores y operadores discursivos en un *corpus* de comentarios escritos de YouTube, desde una perspectiva *corpus-driven* que combina análisis manual y automático. A partir de 1.058 comentarios a videos noticiosos de canales chilenos, se identificaron múltiples expresiones con función de conector u operador que no figuran en los repertorios tradicionales, lo que sugiere una variación activa en el español digital.

Uno de los principales hallazgos es la alta presencia de construcciones semilibres, es decir, unidades parcialmente fijadas que presentan una parte invariable y otra variable, pero que comparten una función discursiva reconocible. Expresiones como *más que*, *oye*, *mira* o *a ver* no responden a la idea de que los conectores y operadores sean necesariamente formas invariables y gramaticalizadas, sino que los sitúa en un continuo formal y funcional que va desde formas plenamente fijadas hasta construcciones emergentes. Desde una perspectiva funcional, destaca la polifuncionalidad y sensibilidad contextual de muchas de estas expresiones, lo que refuerza la pertinencia del enfoque construccional adoptado. Este permite sistematizar no solo unidades fijadas, sino también patrones emergentes con significado procedural, que son fundamentales para la cohesión y la gestión de la subjetividad en la interacción digital.

Otro hallazgo relevante es que las funciones discursivas que presentan mayor diversidad formal –esto es, mayor número de formas distintas utilizadas para expresarlas– coinciden con aquellas que resultaron ser las más productivas en términos de

paradigmas formal-funcionales, es decir, con más variantes formales y mayor frecuencia de uso en el *corpus*. Se trata, en ambos casos, de los conectores ordenadores interactivos, los operadores modales y los operadores argumentativos. Esta convergencia entre diversidad formal y productividad formal-funcional de los paradigmas propuestos confirma que dichas funciones son especialmente activas en los comentarios digitales de YouTube, donde los usuarios no solo exponen sus opiniones, sino que también interactúan con otros comentaristas y con los contenidos de los videos. En este contexto discursivo, marcado por la interactividad, la expresividad y la argumentación, estas funciones permiten organizar el intercambio, expresar valoraciones y sostener o refutar posiciones de manera explícita y flexible.

En consecuencia, estos resultados refuerzan la pertinencia del enfoque construccional adoptado, que no solo permite describir unidades fijadas, sino también sistematizar patrones emergentes de variación formal que reflejan dinámicas propias del discurso digital. Los resultados sugieren que ciertos géneros digitales no solo favorecen el uso intensivo de conectores y operadores, sino que también propician la emergencia de variantes formales novedosas y patrones expresivos menos fijados.

Metodológicamente, la propuesta de paradigmas formal-funcionales permite organizar variantes que comparten función discursiva y parte invariable, facilitando su identificación en grandes volúmenes de datos. Esta estrategia, sumada al procedimiento KWIC20 para la extracción contextual, ofrece herramientas replicables para el análisis de la variación en discursos digitales, con aplicaciones potenciales en lingüística computacional, enseñanza del español y desarrollo de repertorios descriptivos.

Aunque el estudio no aborda directamente la dimensión dialectal, la concentración en el español chileno puede haber influido en la emergencia de ciertas expresiones idiosincráticas. Este aspecto merece una exploración contrastiva futura que permita evaluar su incidencia en los procesos de fijación formal.

En relación con estudios anteriores, los resultados coinciden con investigaciones que evidencian la existencia de unidades en distintas etapas de gramaticalización y lexicalización (San Martín, 2011; Repede, 2023), y aportan un marco sistemático para su identificación. Asimismo, se alinean con líneas recientes que interpretan estas expresiones como construcciones en proceso de fijación (Fuentes Rodríguez, 2020; Narrog & Heine, 2021), proporcionando evidencia empírica desde un *corpus* de interacciones digitales que permite observar estos procesos en tiempo real.

Este trabajo también amplía las observaciones sobre el uso expresivo, emergente y polifuncional de conectores y operadores en redes sociales, aportando una mirada de conjunto que va más allá de estudios centrados en casos individuales. Así, contribuye al análisis del español digital como laboratorio lingüístico, especialmente fértil para examinar fenómenos de cambio y variación en su fase emergente. En este sentido, los hallazgos muestran que la variación en comentarios digitales no solo no es desordenada, sino más dinámica y productiva que aquella observada en contextos de escritura formal (Robledo & Aguilar, 2023, aceptado, Robledo & Rodríguez, 2024, en evaluación), lo que confirma que estos espacios constituyen un terreno privilegiado para observar la evolución discursiva del español.

Finalmente, los resultados pueden contribuir al modelamiento de géneros digitales a partir de los rasgos discursivos de los textos (Nazar, 2024) y se pueden aplicar a la enseñanza del español,

la lexicografía y el desarrollo de tecnologías lingüísticas (Nazar & Renau, 2023; Nazar *et al.*, 2024; Venegas *et al.*, 2024), aportando un marco empírico útil para el análisis de la cohesión y la coherencia en contextos digitales masivos (Herring *et al.*, 2013).

AGRADECIMIENTOS

Agradecemos a la Agencia Nacional de Investigación y Desarrollo del Gobierno de Chile por financiar el Proyecto Fondecyt de Postdoctorado n.º 3230617, “Variación formal de marcadores del discurso en español: extracción y clasificación de variantes mediante técnicas semiautomáticas de explotación de *corpus*”, del cual esta investigación forma parte.

Referencias

- Albelda Marco, Marta. 2016. La expresión de la evidencialidad en la construcción se ve (que). *Spanish in Context* 13(2). 237-262.
- Albelda Marco, Marta. 2018. ¿Atenuación del compromiso del hablante?: el caso de los evidenciales “por lo visto” y “se ve que”. *RILCE* 34(3). 1179-214.
- Androutsopoulos, Jannis. 2013. Code-switching in computer-mediated communication. En Susan C. Herring, Dieter Stein & Tuija Virtanen (eds.), *Pragmatics of computer-mediated communication*, 667-694. Berlín/Boston: De Gruyter.
- Androutsopoulos, Jannis. 2015. Networked multilingualism: Some language practices on Facebook and their implications. *International Journal of Bilingualism* 19(2). 185-205.
- Artstein, Ron & Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics* 34(4). 555-596.
- Aulit, Laetitia. 2021. La atenuación en interacciones controvertidas en comentarios sobre noticias digitales y en Twitter. *Normas* 11(1). 1-18.
- Baker, Paul. 2023. *Using corpora in discourse analysis*. 2nd ed. Londres: Bloomsbury.
- Baroni, Marco, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation* 43. 209-226.
- Bayerl, Petra Saskia & Sabine Annette Paul. 2011. What determines inter-coder agreement in manual annotations? A meta-analytic investigation. *Computational Linguistics* 37(4). 699-725.
- Bestgen, Yves. 2024. Back to basics in measuring lexical diversity: Too simple to be true. *Applied Linguistics* 45. 926-932.
- Borreguero Zuloaga, Margarita Natalia. 2017. Los relatos coloquiales: partículas discursivas y polifonía. *Pragmalingüística* 25. 62-88.
- Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.
- Briz, Antonio. 2014. Hablar electrónicamente por escrito. *CHIMERA* 1. 77-89.
- Briz, Antonio, Salvador Pons & José Portolés (coords.). 2008. *Diccionario de partículas discursivas del español* [en línea]. Disponible en: <http://www.dpde.es>

- Cantamutto, Lucía & Cristina Vela Delfa. 2020. Mensajes, publicaciones, comentarios y otros textos breves de la comunicación digital. *Tonos Digital* 38, 1-27.
- Chen, Xueqing. 2022. *El análisis del uso y función de la ironía en Twitter sobre el COVID-19*. Tesis de maestría, Universitat Politècnica de València.
- Crystal, David. 2008. *Txtng: The gr8 db8*. Oxford: Oxford University Press.
- De Luca, Natalia. 2021. Análisis del marcador “tipo” en memes: hacia el concepto de marcador-meme en WhatsApp de interacciones juveniles. *Revista Estudios del Discurso Digital (REDD)* (4). 49-70.
- De Piero, José Luis & María Evangelina Narvaja. 2018. Representaciones de género en comentarios digitales en dos Fan Page de Facebook. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud* 16(2). 759-775.
- Degand, Liesbeth & Anne-Marie Simon-Vandenbergen. 2011. Grammaticalization and (inter)subjectification of discourse markers. *Linguistics* 49(2). 287-294.
- Deng, Delin. 2023. The grammaticalization of the discourse marker genre in Swiss French. *Languages* 8(1). 28.
- Elvira, Javier. 2015. *Lingüística histórica y cambio gramatical*. Madrid: Síntesis.
- Espinal, M. Teresa, Cristina Real-Puigdollers & Xavier Villalba. 2024. From a movement verb to an epistemic discourse marker: The diachronic change of Spanish *vaya*. *Linguistic Variation* 24(1). 1-36.
- Falusí, Kinga. 2021. El marcador metadiscursivo de control de contacto ¿cierto? ¿un fenómeno hispanoamericano? *Nueva Revista de Filología Hispánica* 69(1). 43-77.
- Fuentes Rodríguez, Catalina. 2003. Operador/conector, un criterio para la sintaxis discursiva. *RILCE* 19(1). 61-85.
- Fuentes Rodríguez, Catalina. 2018. *Diccionario de conectores y operadores del español*. 2a ed. Madrid: Arco/Libros.
- Fuentes Rodríguez, Catalina. 2020a. Categorías, relaciones causativas y argumentativas: de ahí (que)/ de hecho. En A. Messias, C. Fuentes & M. Martí (eds.), *Aportaciones desde el español y el portugués a los marcadores discursivos: treinta años después de Martín Zorraquino y Portolés*, 28-47. Sevilla: Universidad de Sevilla.
- Fuentes Rodríguez, Catalina. 2020b. Mira/mira que: construcciones, contexto argumentativo y funciones relacionales. *RILCE* 36(3). 941-966.
- Fuentes Rodríguez, Catalina. 2024a. *Macrosintaxis del español*. Berlín/Boston: De Gruyter.
- Fuentes Rodríguez, Catalina. 2024b. Macrosintaxis de la construcción «no (te) creas». *RILCE* 40(3). 903-936.
- Fuentes Rodríguez, Catalina. 2025. Géneros discursivos y criterios para su clasificación. En M. E. Placencia & A. Parini (eds.), *Introducción al estudio del discurso digital en español*, 7-24. Londres: Routledge.

- Fuentes Rodríguez, Catalina & Marina González-Sanz. 2024. *La enunciación en el discurso*. Madrid: Arco/Libros.
- Garachana Camarero, Mar. 2024. La creación de marcadores discursivos en español. En S.N. Dworkin, G. Clavería & Á.S.O. De Toledo y Huerta (eds.), *Lingüística histórica del español: The Routledge Handbook of Spanish Historical Linguistics*, 111-122. Londres: Routledge.
- Giumetti, Gary W. & Robin M. Kowalski. 2022. Cyberbullying via Social Media and Well-Being. *Current Opinion in Psychology* 45(8). 101314.
- Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Grande Alija, Francisco Javier. 2020. Para que luego digan...: de la sintaxis oracional al discurso. *RILCE* 36(3).1045-1072.
- Grande Alija, Francisco Javier. 2024. Subordinación y coordinación: de la sintaxis al discurso. *Círculo de Lingüística Aplicada a la Comunicación* 100. 75-87.
- Guillén Escamilla, Josaphat Enrique. 2022. Haz de cuenta (que) como marcador discursivo del español de México. Un estudio de variación pragmática a partir del análisis de corpus. *Cuadernos de Lingüística de El Colegio de México* 9. 1-45.
- Guillén Escamilla, Josaphat Enrique. 2023. Estudio en tiempo real del marcador discursivo «ya te/le digo» en la variante del español de la Ciudad de México. *Forma y Función* 36(1). e6951.
- Guillén Escamilla, Josaphat Enrique. 2024. Así de (que) como introductor de estilo directo en el español de México. Un estudio a partir del análisis de corpus. *Oralia* 27(1). 31-48.
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6). 1205-1247.
- Herring, Susan C. & Ashley Dainas. 2017. “Nice picture comment!” . Graphicons in Facebook comment threads. *Proceedings of the 50th Hawaii international conference on system sciences*, 2185-2194.
- Herring, Susan C., Dieter Stein & Tuija Virtanen. 2013. *Pragmatics of Computer-Mediated Communication*. Berlín: De Gruyter.
- Hoffmann Hölcéll, Brenda Denise. 2023. *Todo lo que es: usos y construccionalización de un mecanismo focal*. Tesis de doctorado, Pontificia Universidad Católica Argentina.
- Jansegers, Marlies & Marta Albelda Marco. 2020. Estudio contrastivo de los evidenciales “por lo visto” y “se ve (que)” en las variantes dialectales del español. En Antonio Messias, Catalina Fuentes & Manuel Martí (coords.), *Aportaciones desde el español y el portugués a los marcadores discursivos: treinta años después de Martín Zorraquino y Portolés*, 363-384. Sevilla: Universidad de Sevilla.
- Kozhukhova, Irina & Oxana Pastukhova. 2019. Discourse markers and their functions in new digital media (Based on Buzzfeednews). En Z. Marina Viktorovna (ed.), *Journalistic text in a new technological environment: Achievements and problems*, 64-71. Future Academy.

- Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33. 159-174.
- López Quero, Salvador. 2018. La función apelativa del marcador discursivo venga en Twitter. *Onomázein* 39. 48-65.
- López Quero, Salvador. 2019. El marcador discursivo ahora con verba dicendi en Twitter. *Círculo de Lingüística Aplicada a la Comunicación* 77. 243-260.
- Loureda, Óscar & Esperanza Acín. 2010. *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco/Libros.
- Malvern, David, Brian Richards, Ngoni Chipere & Pilar Durán. 2004. *Lexical diversity and language development: Quantification and assessment*. Palgrave Macmillan.
- Martín Zorraquino, María Antonia & José Portolés. 1999. Los marcadores del discurso. En Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. 3, 4051-4213. Madrid: Espasa Calpe.
- Maruenda-Bataller, Sergio, Laura Mercé & Elena Castellano-Ortolà (eds.). 2024. *Discourse, gender, and violence: Insights from news and social media texts*. Peter Lang.
- Montolío, Estrella. 2001. *Conectores de la lengua escrita. Contraargumentativos, consecutivos, aditivos y organizadores de la información*. Barcelona: Ariel.
- Montolío, Estrella. 2013. Construcciones conectivas que encapsulan. [A pesar de + SN] y la escritura experta. *Cuadernos AISPI* 2. 115-131.
- Narrog, Heiko & Bernd Heine. 2021. *Grammaticalization*. Oxford: Oxford University Press.
- Nazar, Rogelio. 2024. Statistical modeling of discourse genres: The case of the opinion column in Spanish. *SN Computer Science* 5. 959.
- Nazar, Rogelio & Irene Renau. 2023. Estilector: un sistema de evaluación automática de la escritura académica en castellano. *Perspectiva Educacional* 62(2). 37-59.
- Nazar, Rogelio, Irene Renau & Hernán Robledo. 2024. Dismark and Text·a·Gram: Automatic identification and categorization of discourse markers in texts. En Cecilia-Mihaela Popescu & Oana-Adriana Dută (eds.), *Discourse markers in Romance languages: Crosslinguistic approaches in Romance and beyond*, 165-184. Berlín: Peter Lang.
- Onodera, Noriko. 2011. The grammaticalization of discourse markers. En *The Oxford handbook of grammaticalization*, 614-624. Oxford: Oxford University Press.
- Padilla Herrada, María Soledad. 2023. Los operadores discursivos de lujo y del tirón en el discurso digital. *Círculo de Lingüística Aplicada a la Comunicación* (93). 259-269.
- Padilla Herrada, María Soledad. 2024. La afirmación y los marcadores discursivos: una propuesta terminológica de los contenidos afirmativos interactivos. *Logos* 34(1). 167-188.
- Pano Alamán, Ana. 2020. Marcadores discursivos y español coloquial en las redes sociales. *Doxa Comunicación* 31. 381-401.

- Parini, Alejandro, Edgardo Galende & Verónica Vera. 2025. YouTube. En María Elena Placencia & Alejandro Parini (eds.), *Introducción al estudio del discurso digital en español*, 67-80. Routledge.
- Pavón Lucero, María Victoria. 2003. *Sintaxis de las partículas*. Madrid: Visor.
- Pérez-Sabater, Carmen. 2024. *La informalización del discurso digital público: reflexiones sobre el inglés y el español*. Londres/Nueva York: Routledge.
- Pérez-Sabater, Carmen & Ginette Maguelouk-Moffo. 2019. Managing identity in football communities on Facebook: Language preference and language mixing strategies. *Lingua* 225. 32-49.
- Placencia, María Elena & Alejandro Parini (eds.). 2025. *Introducción al estudio del discurso digital en español*. Londres/Nueva York: Routledge.
- Pomikálek, Jan. 2011. *Removing boilerplate and duplicate content from web corpora*. Tesis de doctorado, Masaryk University.
- Pons Bordería, Salvador. 2006. A functional approach for the study of discourse markers. En Kerstin Fischer (ed.), *Approaches to discourse particles*, 77-99. Elsevier.
- Pons Bordería, Salvador. 2013. Un solo tipo de reformulación. *Cuadernos AISPI* (2). 151-170.
- Portolés, José. 2001. *Marcadores del discurso*. Barcelona: Ariel.
- Posteguillo, Santiago. 2003. *Netlinguistics: Language, Discourse and Ideology in Internet*. Castelló de la Plana: Universitat Jaume I.
- Ray, Geraldine, Christopher D. McDermott & Mathew Nicho. 2024. Cyberbullying on social media: Definitions, prevalence, and impact challenges. *Journal of Cybersecurity* 10(1). tyae026.
- Rayson, Paul. 2008. From key words to key semantic domains. *International Journal of Corpus Linguistics* 13(4). 519-549.
- Repede, Doina. 2023. Estudio piloto sobre los marcadores del discurso como introductores de cita directa en el Corpus oral PERUSEV. *Lingüística* 39(2). 11-32.
- Robledo, Hernán. 2021. *Categorización de los marcadores del discurso del español: una propuesta inductiva guiada por corpus paralelo*. Tesis de doctorado, Pontificia Universidad Católica de Valparaíso.
- Robledo, Hernán & Carlos Aguilar. 2023. Variación formal de los marcadores del discurso del español: estudio de un corpus de resúmenes de artículos de investigación lingüística. Ponencia presentada en el *XXIII Congreso Internacional de la Sociedad Chilena de Lingüística*. Santiago, Chile.
- Robledo, Hernán & Carlos Aguilar. aceptado. Variantes formales de marcadores del discurso en español: exploración en resúmenes de artículos de investigación lingüística. *Revista Literatura y Lingüística*.
- Robledo, Hernán & Rogelio Nazar. 2018. Clasificación automatizada de marcadores discursivos. *Procesamiento del Lenguaje Natural* 61. 109-116.

- Robledo, Hernán & Rogelio Nazar. 2023. A proposal for the inductive categorisation of parenthetical discourse markers in Spanish using parallel corpora. *International Journal of Corpus Linguistics* 28(4). 500-527.
- Robledo, Hernán & Antonia Rodríguez. 2024. Variación formal de los marcadores del discurso: análisis de un corpus de introducciones de memorias de titulación de Ingeniería Mecánica. Ponencia presentada en el *XX Congreso Internacional de la Asociación de Lingüística y Filología de América Latina*. Concepción, Chile.
- Robledo, Hernán & Antonia Rodríguez. en evaluación. Variación formal de los marcadores del discurso del español: exploración de un corpus de memorias de titulación. *Nueva Revista del Pacífico*.
- Rodríguez-Abruñeras, Paula. 2020. Outlining a grammaticalization path for the Spanish formula en plan (de): A contribution to crosslinguistic pragmatics. *Linguistics* 58(6). 1543-1579.
- Sampietro, Agnese. 2016. *Emoticonos y emojis: análisis de su historia, difusión y uso en la comunicación digital actual*. Tesis de doctorado, Universitat de València.
- San Martín, Abelardo. 2011. Los marcadores interrogativos de control de contacto en el corpus PRESEEA de Santiago de Chile. *Boletín de Filología* 46(2). 135-166.
- San Martín, Abelardo. 2015. Variantes y equivalentes funcionales de al final: los reformuladores de recapitulación en el habla santiaguina. *RLA. Revista de lingüística teórica y aplicada* 53(2). 97-119.
- Sandoval, Jorge & Abelardo San Martín. 2020. Los modalizadores de atenuación como (que), igual, medio/a, de repente y capaz (que) en el habla santiaguina: análisis pragmático y sociolingüístico. *Nueva Revista del Pacífico* (72). 144-172.
- Scheinbaum, Angeline Close (ed.). 2024. *The darker side of social media: Consumer psychology and mental health*. 2nd ed. Routledge.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Scott, Mike. 2010. *WordSmith tools manual version 5.0*. Liverpool: Lexical Analysis Software.
- Suárez Hernández, Ariana. 2023. Construcciones formadas sobre la base duda: una perspectiva diacrónica. *Dicenda* 41. 213-223.
- Tannen, Deborah & Anna Marie Trester (eds.). 2013. *Discourse 2.0: Language and new media*. Washington, DC: Georgetown University Press.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Ámsterdam: John Benjamins.
- Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- Tweedie, Fiona J. & R. Harald Baayen. 1998. How variable may a constant be? Measures of lexical richness in perspective. *Computers and the Humanities* 32. 323-352.
- Uceda Leal, María. 2023. Un enfoque sociopragmático para el análisis de *dilo y preguntó* como operadores discursivos gestados en Twitter. *Biblioteca de Babel* 4. 149-171.

- Vande Castele, An & Catalina Fuentes Rodríguez. 2019. La multifuncionalidad en la traducción del marcador discursivo pues. *ELUA* (Anexo VI). 201-216.
- Vandergriff, Ilona. 2013. Emotive communication online: A contextual analysis of computer-mediated communication (CMC) cues. *Journal of Pragmatics* 51. 1-12.
- Vela Delfa, Cristina & Lucía Cantamutto. 2021. *Los emojis en la interacción digital escrita*. Madrid: Arco/Libros.
- Venegas, René, Eilene Martha Castro Cano & Dyicela Cornejo. 2024. Marcadores discursivos y calidad de la escritura: estudio empírico-computacional del ensayo en estudiantes universitarios. *Íkala* 29(2). 1-21.
- Webber, Bonnie, Rashmi Prasad, Alan Lee & Aravind Joshi. 2019. *The Penn Discourse Treebank 3.0 Annotation Manual*. University of Pennsylvania, Linguistic Data Consortium.
- Wiśniewska, Natalia. 2021. Operadores argumentativos como mecanismo lingüístico de la argumentación utilizado en Instagram: el caso de las cuentas #studygram. *Tonos Digital* (40). 1-22.
- Yus, Francisco. 2010. *Ciberpragmática 2.0: nuevos usos del lenguaje en Internet*. Barcelona: Ariel.
- Yus, Francisco. 2014. Not all emoticons are created equal. *Linguagem em Discurso* 14(3). 511-529.
- Yus, Francisco. 2022. *Smartphone communication: Interactions in the app ecosystem*. Londres/Nueva York: Routledge.
- Zambrano Ojeda, Isis Nataly. 2023. ¡Achachay, qué frío, ve! Subjetivización de ‘ve’ como un marcador discursivo. *Cuadernos de Lingüística de El Colegio de México* 10. e266.