

¿SE INVESTIGA O INVESTIGAMOS?: USOS DE LA MODALIZACIÓN EN ARTÍCULOS DE INVESTIGACIÓN RELACIONADOS CON EL LENGUAJE

ANDREA TORRES PERDIGÓN

Pontificia Universidad Javeriana (Colombia)
ORCID: <https://orcid.org/0000-0003-2539-4766>

DIANA MORENO RODRÍGUEZ

Pontificia Universidad Javeriana (Colombia)
ORCID: <https://orcid.org/0000-0002-1633-5265>

NATHALIA LAMPREA ABRIL

Pontificia Universidad Javeriana (Colombia)
ORCID: <https://orcid.org/0000-0002-7064-9273>

RESUMEN: Este texto expone los resultados preliminares del análisis de un corpus de 1260 artículos publicados en español entre 2002 y 2018 relacionados con el tema del lenguaje. Los artículos fueron publicados en revistas latinoamericanas de ciencias sociales y humanidades, y fueron tomados del índice citacional de SciELO. El propósito es establecer marcas frecuentes de la modalización en el corpus, para luego analizar su funcionamiento con respecto a aquello que llamamos efectos de despersonalización o de personalización de este género discursivo. El estudio se realizó a partir de la lingüística de corpus y de la categoría de modalización, y se llevó a cabo por medio de dos programas: Iramuteq y Lexico 5. Los resultados muestran la dominancia de marcas delocutivas en los segmentos repetidos (en especial en categorías de generalización, condición, oposición y posibilidad). Estos resultados implican que las marcas delocutivas se formulan en expresiones relativamente fijas y que parecen vincularse con el posicionamiento de los autores, posicionamiento que se comunica por medio de efectos de impersonalidad.

PALABRAS CLAVE: Lingüística de corpus, artículo, géneros discursivos, modalización, segmentos repetidos

*ONE INVESTIGATES OR WE INVESTIGATE? USES OF MODALIZATION IN RESEARCH
ARTICLES RELATED TO LANGUAGE*

ABSTRACT: This text exposes the first results of an analysis of a corpus composed by 1 260 articles about language published in Spanish between 2002 and 2018. The articles were published in Latin American specialized journals in social sciences and humanities and were downloaded from the citational index SciELO. The purpose of this article is to establish frequent marks of modalization within the corpus, in order to analyze its use and function regarding what we call impersonality and personality effects of this discursive genre. The study was developed following the principles of corpus linguistics as well as the modalization category. Two software programs were used: Iramuteq and Lexico 5. Results show the prevailing frequency of delocutive marks in the repeated segments, particularly in those that we classified as generalization, condition, opposition and possibility. These results indicate that delocutive marks are enunciated as relatively fixed expressions and they seem to relate to author's positions, positions which, as we argue, seem to be communicated by means of an impersonality effect.

KEYWORDS: Corpus linguistics, article, discursive genres, modalization, repeated segments

1. INTRODUCCIÓN

Precisar qué significa investigar puede ser problemático hoy. No tanto porque determinadas comunidades discursivas y académicas no lo tengan claro, sino porque, precisamente, las disciplinas y las ciencias establecen criterios para sus prácticas investigativas que no son uniformes (y que, quizá, no tendrían por qué serlo). En otras palabras, lo que se entiende y valora como investigación puede ser heterogéneo y depende, en gran medida, de las áreas de saber y de las disciplinas, así como de momentos históricos determinados. Si bien la originalidad y el aporte al campo de estudio parecerían características de lo que consideramos investigación hoy en día, las marcas lingüísticas y textuales que revelan esa novedad no son las mismas para la biología, la filosofía o la lingüística, por poner solo algunos ejemplos.

Sin embargo, a pesar de esa heterogeneidad, hay unos tipos de textos que parecen ser comunes como forma de producción y divulgación de la investigación: en campos del saber muy variados circulan artículos de investigación y libros colectivos o individuales de carácter investigativo. Esta forma de circulación implica que hay una superficie textual y unos géneros discursivos compartidos, aspecto particularmente marcado en las ciencias sociales y en las humanidades, en las que la comunicación de lo investigado se establece en géneros primordialmente escritos –aunque no de forma exclusiva–.

En este orden de ideas, en este artículo presentamos los resultados parciales de un estudio¹ llevado a cabo entre 2018 y 2019. Globalmente, el estudio pretende rastrear

¹ El nombre del proyecto fue *Concepciones de la investigación en estudios relacionados con el lenguaje* y contó con cuatro investigadores: Andrea Torres Perdigón, Diana Moreno Rodríguez, Nathalia Lamprea Abril y Diana Lucio-Arias.

características dominantes de este género discursivo, específicamente en un corpus de 1260 artículos que tratan el problema del lenguaje y que han sido publicados en español durante las últimas dos décadas (2002 a 2018) en revistas latinoamericanas. El corpus se construyó a partir del tema del lenguaje porque nos interesaba rastrear cómo distintas disciplinas se han aproximado al tema y ver su vigencia en la región latinoamericana en los últimos años².

En un principio nuestro estudio se interesó por cómo se evidencian las concepciones de la investigación en artículos que provienen de disciplinas que estudian o trabajan con ciertas nociones de lenguaje, y puntualmente por: ¿qué rasgos o marcas lingüísticas y textuales pueden dar cuenta de cómo se posicionan los sujetos de enunciación ante su discurso? De una manera más específica, este artículo procura establecer relaciones entre la materialidad de los textos en los cuales circula la investigación a través de un género discursivo concreto –el artículo– y posibles rasgos de concepciones de investigación contemporáneas. Se trata de una pregunta amplia que hemos puesto en diálogo con el enfoque de la lingüística de corpus (Mayaffre, 2005; Rastier, 2011; Tognini-Bonelli, 2001) por dos razones. La primera es que nos permite un movimiento constante entre los postulados teóricos y el análisis empírico de un conjunto de textos concretos (Rastier, 2011). La oscilación entre teorías lingüísticas y el análisis de un corpus amplio de documentos nos permite pensar en nuevos observables y en las manifestaciones lingüísticas y textuales específicas que podrían relacionarse en un futuro con ciertas concepciones de investigación; así evitamos en lo posible la imposición o el forzamiento de un marco teórico enteramente predeterminado al corpus de artículos. La segunda razón es que el propósito consiste en estudiar la materialidad de los textos en los que la investigación se produce, circula y se comunica en ciencias sociales y humanas, y su posible relación con cómo se concibe la investigación. En esta medida, el término de “concepción” no se entiende aquí como ligado a las percepciones o a las representaciones mentales y sociales (Moscovici, 1989) de los investigadores o de los académicos, sino a aquello que produce unos efectos en textos concretos dentro de lo que distintas disciplinas consideran válido como “artículo de investigación”. Tampoco nos interesa determinar –de manera ingenua o normativa– si el artículo es un género objetivo y transparente o si, por el contrario, es argumentativo y persuasivo: la indagación sobre el género del artículo va un poco más allá porque busca acercarse a comprender algunos rasgos de su funcionamiento a través de un corpus en español bastante grande que cubre casi dos décadas. Dado que este estudio se centra en textos singulares (y no en concepciones abstractas o mentales), la perspectiva de la lingüística de corpus nos permitió un análisis del género discursivo –tal y como lo plantea inicialmente Mijaíl Bajtín (1999)– de lo que identificamos como artículo de investigación en una escala mucho más amplia de lo que permiten la lectura y el análisis individual de los textos, sobre todo sin asistencia de herramientas computacionales.

² Los artículos fueron publicados en revistas que se clasifican como de lingüística, filosofía, estudios de mujeres, comunicación, literatura, antropología, arte, sociología, estudios culturales, arqueología, historia, estudios étnicos, religión, educación y ciencias sociales.

Por esa razón, el tipo de análisis que proponemos emplea la lingüística de corpus como un enfoque metodológico. Sin embargo, esto no se debe a un ánimo descriptivo frente a la lengua española contemporánea, sino a una voluntad de identificar qué aspectos lexicométricos y nuevos observables del corpus pueden estar relacionados con determinados efectos de la escritura de investigación en Hispanoamérica. En esta medida, el foco de interés no es tanto describir el funcionamiento de la lengua española en sus variedades diatópicas hispanoamericanas, sino más bien observar qué marcas lingüísticas y textuales recurrentes podrían vincularse con algunos rasgos de cómo se concibe la investigación en distintas áreas.

Ahora bien, dentro de los múltiples rasgos que podrían relacionarse con concepciones de investigación diversas en un corpus de artículos, decidimos centrarnos inicialmente en uno. Al hacer la primera lectura y revisión de los resúmenes del corpus final³, el equipo de investigadoras empezó a manifestar una primera intuición: contra las expectativas genéricas, las posturas de los autores eran visibles o invisibles de maneras muy heterogéneas. La invisibilización podría estar ligada, en un primer momento, a cierto efecto de despersonalización en la enunciación y a la manera de modalizar el discurso. A raíz de esto, el fenómeno de modalización (Charaudeau, 1992) se eligió como un primer elemento para estudiar las relaciones entre lo descriptivo, lo explicativo, lo propositivo y lo argumentativo en este corpus, puesto que permitiría reconocer marcas concretas de posturas autorales en los textos.

El propósito de este texto es, entonces, establecer marcas frecuentes de la modalización en el corpus de artículos de investigación para luego analizar su uso y su funcionamiento con respecto a aquello que llamaremos efectos de personalización o de despersonalización de este género discursivo, efectos que darían cuenta de cómo el sujeto se construye y construye su discurso. La pregunta que orienta este artículo es la siguiente: ¿qué marcas lingüísticas y textuales indican una modalización de los sujetos de la enunciación dentro del corpus de artículos y cómo operan? Para desarrollar nuestro propósito y responder al menos parcialmente a esta pregunta, presentaremos de manera breve el estado de la cuestión, luego nuestra perspectiva teórica para después explicar la metodología que permitió hacer ese primer análisis y, finalmente, hacer énfasis en los resultados preliminares.

³ Después de realizar la búsqueda de artículos en el índice citacional de SciELO, se construyó una base de datos con la información bibliográfica de los artículos descargados. Luego, se seleccionó un 5 % de los resúmenes de manera aleatoria y todos los miembros del equipo hicieron una primera revisión de la pertinencia de esta muestra. A partir de esa primera lectura, se establecieron cinco preguntas orientadoras que permitieron llegar a un porcentaje de concordancia de 73 %. Estas preguntas fueron usadas para leer todos los resúmenes de la base de datos que fueron distribuidos equitativamente entre cuatro evaluadores (25 % cada uno). Los casos dudosos de la pertinencia de algún resumen se resolvían en reuniones. Esta fue la manera de depurar la base de datos y, posteriormente, descargar el corpus completo. Las preguntas orientadoras fueron las siguientes: 1. ¿El artículo da cuenta de una concepción de investigación? 2. ¿El artículo aborda el lenguaje? 3. ¿Es el lenguaje un elemento central o accesorio en el ejercicio investigativo? 4. ¿Se da cuenta de una preocupación por el lenguaje? 5. ¿El problema de investigación está relacionado con el lenguaje?

2. ESTADO DE LA CUESTIÓN

La inquietud por las características del discurso científico y académico ha sido estudiada desde distintas perspectivas en los últimos veinte años. En especial dentro del ámbito de los estudios del lenguaje, la cuestión que quizá empieza a despertar el interés por este problema tiene que ver con el texto fundador de Bajtín en cuanto a los géneros discursivos. En su libro *Estética de la creación verbal*, Bajtín propone un enfoque comunicativo para entender cómo funcionan los enunciados (y por qué se diferencian de las meras oraciones) y cómo los géneros son cadenas de enunciados relativamente estables que comparten temas, estructuras composicionales y estilos (Bajtín, 1999), y cómo se producen en diversas esferas sociales. A partir de esta propuesta, los géneros empiezan a estudiarse ya no como un asunto exclusivo de la teoría literaria ni en función solo de géneros literarios, sino como parte de una interacción entre interlocutores: en términos de Bajtín, pasan a entenderse como géneros discursivos por el carácter dialógico y de interlocución que implican, y esto incluso en los géneros secundarios (primordialmente escritos) que resultan más complejos. El artículo científico, tanto en humanidades como en ciencias naturales o sociales, empezó a estudiarse desde esta perspectiva, más allá de un enfoque prescriptivo de cómo escribir bien en géneros determinados⁴. Dentro de las investigaciones que se han acercado al estudio del artículo científico como género existen, en particular en la tradición anglosajona, aproximaciones desde la lingüística aplicada, como las de Swales (1990, 2004) y desde los estudios sobre literacidad y escritura académica (Bazerman y otros, 2009). Este tipo de enfoques se ha desarrollado principalmente en lengua inglesa.

En el ámbito hispanohablante, encontramos dos aproximaciones al estudio de los artículos de investigación definidas: una que parte de la inquietud por las manifestaciones de la subjetividad en estos textos, por un lado, y otra que construye su análisis desde estudios de literacidad y enseñanza de escritura académica, por otro. En el primer caso tenemos estudios que se acercan a un enfoque del análisis del discurso, mientras que en el segundo el género del artículo se estudia principalmente con fines pedagógicos, es decir, para enseñar escritura académica en las universidades y en las diferentes disciplinas y campos del saber.

En la primera aproximación se encuentran trabajos como el de García Negroni (2008), quien se ha interesado por abordar el problema de la subjetividad y de la polifonía en artículos de investigación en distintas áreas. El propósito de esta autora es mostrar que la visión tradicional del artículo científico, como un texto objetivo e impersonal, ha sido reevaluada y que la autoría en el “discurso científico-académico (incluso el de las llamadas ‘ciencias duras’) deja huellas de su presencia y que su discurso no es ni monológico ni neutro” (García Negroni, 2008: 11). Para este fin, la autora construye un corpus de artículos (de 190 621 palabras) provenientes de cuatro disciplinas: historia, lingüística, geología y medicina. En ellos analiza dos formas determinadas de cómo

⁴ Ver Moreno y otros (2010); Suárez Relinque y otros (2013); Murillo y otros (2017) o Santesteban-Echarri y Núñez-Morales (2017).

se manifiesta la subjetividad: las marcas de persona y las maneras de referir la palabra ajena. Sus resultados consisten en rastrear y construir un inventario de casos del uso de las marcas de despersonalización y personalización. En el primer caso, encuentra el uso de nominalizaciones como “el análisis expuesto en este trabajo se propone abrir” (2008:12), metonimias como “Esta contribución analiza” (2008:12), estructuras impersonales con infinitivo como “cabe destacar” y el uso de pasivas con “ser” y “se” como “Más arriba se planteó” o “la cual en parte ha sido tomada” (2008:13). En el segundo caso, encuentra el uso de la primera persona plural (de referencia genérica, un “nosotros” inclusivo con el lector y otro coincidente) y la primera persona singular coincidente. En cuanto a las maneras de referir la palabra ajena, García Negroni muestra cinco modalidades: referencias no integradas (en las que solo se mencionan los resultados), referencias parcialmente integradas (indicadas bibliográficamente), referencias semiintegradas (citas de parafraseo con el nombre del autor vinculado a la oración), citas integradas (es decir, textuales incorporadas al enunciado de los autores) y citas destacadas (más extensas y separadas del cuerpo del texto). La autora encuentra que en las cuatro disciplinas que abarca su estudio, y a pesar del carácter limitado de su corpus, es posible afirmar la dimensión subjetiva y polifónica del género del artículo de investigación. Además, establece que las formas de subjetividad y polifonía muestran la configuración de *ethos* diferentes según las disciplinas, en particular en cuanto a cuáles se considerarían más neutras y asépticas o cuáles más atentas al propio discurso. Sin embargo, aparte de ese *ethos* disciplinar, para ella las manifestaciones de subjetividad son las que permiten “[...] que el locutor-autor pued[a] posicionarse en relación con la tradición científica en la que busca inscribir su investigación y promocionar, en concomitancia, el valor de su trabajo” (García Negroni, 2008: 28). Otros estudios que se han realizado en esta dirección son Bolívar (2004, 2005) y García Negroni (2005) y García Negroni y Tordesillas (2001), García Negroni y Ramírez Gelbes (2005) y García Negroni, Estrada y Ramírez Gelbes (2006).

En la segunda aproximación al estudio de los artículos de investigación en español encontramos el enfoque desde los estudios de literacidad, ligados a la lingüística aplicada y la enseñanza de lenguas. En este, varias publicaciones de un grupo de profesores en Chile se han centrado en el análisis de corpus tomado de SciELO Chile, para analizar marcas estructurales de los objetivos, las justificaciones y las conclusiones de los artículos científicos. En estos textos, encontramos una aplicación y adaptación a textos escritos en español de la propuesta de las *rethorical moves*, de Swales (1990, 2004), que consiste en “la expresión de un propósito comunicativo que se asocia a un fragmento textual, y que contribuye al logro del propósito global de un género” (Sabaj Meruane y otros, 2011: 247). En Sabaj Meruane y otros (2011), por ejemplo, el objetivo fue construir un modelo de movidas retóricas adaptando el de Swales a un corpus de 24 artículos de investigación tomados de SciELO (Chile). Metodológicamente, los autores procedieron en tres fases para construir su modelo de movidas retóricas para artículos en español: “(1) indagación sistemática de AI [artículos de investigación] especializados en MR [movidas retóricas], (2) realización de un estudio de casos exploratorio, y (3) homologación de la información y construcción del modelo” (Sabaj Meruane y otros, 2011: 251). El análisis de los artículos fue

realizado manualmente por ocho asistentes de investigación, quienes llenaban una rejilla con campos predeterminados que contenían las *rethorical moves* establecidas de antemano. El resultado general en este pequeño corpus es la construcción de una matriz de movidas retóricas que puedan describir distintas secciones de los artículos de investigación. Se trata de descripciones en términos proposicionales, de modo que el problema de la enunciación o de la modalización está ausente.

Otros ejemplos se encuentran en Fuentes Cortés (2012, 2013), González Arias (2011), Sabaj y otros (2013), Sabaj Meruane y Landea Balin (2012). En ellos se procede a partir de las movidas retóricas para establecer particularidades de algunas disciplinas, como la historia, o de las secciones como la introducción y las conclusiones. Un poco al margen de este campo de las movidas retóricas, está también el estudio de Millaray Salas Valdebenito (2015), a partir del modelo reflexivo del metadiscurso de Ädel (2006), de 238 artículos de investigación de tres disciplinas (medicina, economía y lingüística) que se analizaron manualmente. Los resultados de este estudio consisten en una propuesta de taxonomía de unidades metadiscursivas (es decir, en las que se ve la reflexividad del discurso). Globalmente, estos estudios buscan establecer determinadas marcas de la escritura científica con el ánimo de tomar de allí herramientas para la enseñanza de la escritura académica y científica.

Si bien nuestro estudio parte de la inquietud por algunas formas de despersonalización y de personalización dentro del género del artículo científico, este adopta un enfoque diferente al de la subjetividad y la polifonía, y al de la alfabetización científica y académica. En nuestro caso, las preguntas que se le realizan al corpus se centran en los *efectos* que buscan o producen esos textos desde su enunciación (y, dentro de esta, su modalización). En esta medida, el interés no se da tanto en la expresión o la manifestación de la subjetividad de los autores, en la verdad de las afirmaciones sobre las investigaciones o en las características que permitirían enseñar a escribir este tipo de textos. Nuestra indagación nos lleva a pensar más bien en cómo la modalización de quienes enuncian opera en estos textos y, de esa forma, produce determinados efectos en los artículos, efectos que pueden ser personalizadores o despersonalizadores, ya sea en su argumentación o en su exposición. En otros términos, la idea es observar cómo ciertas marcas de modalización podrían estar relacionadas con lo que se concibe como investigativo en distintas ciencias sociales y en las humanidades, específicamente en artículos relacionados con el tema del lenguaje.

Con respecto a la segunda perspectiva, cabe señalar otra diferencia frente a nuestra apuesta: el presente estudio no parte de las construcciones teóricas para aplicarlas a un corpus de estudio –como es el caso de los *Top-bottom approaches*, empleados en investigaciones como las de Sabaj Meruane y otros (2011)–, sino que recurre a la lingüística de corpus como metodología. Lo anterior implica que pone en diálogo constante los postulados teóricos con respecto a la modalización y el análisis del corpus mismo y sus marcas lingüísticas y textuales. También es importante precisar que esta investigación contempla un corpus mucho más amplio que los trabajados en los estudios mencionados y que, en esa medida, es pertinente que se nutra de la lingüística de corpus como enfoque y del tratamiento automatizado de lengua natural

apoyado en la lexicometría/textometría⁵. El corpus de 1260 artículos en español que hemos construido implica que los resultados serán quizá más robustos y con mayor capacidad explicativa que los estudios del actual estado del arte en español.

3. MARCO TEÓRICO

Nuestra orientación teórica y conceptual implica un diálogo entre tres elementos de la lingüística y el análisis del discurso. El primer elemento es la problematización del género discursivo, desde el ya clásico texto de Bajtín publicado en 1979 (1999) pasando por aproximaciones como las de Jean-Michel Adam (1997, 2011b, 2011a). El segundo es el postulado de planos o niveles textuales, como los ha trabajado en particular Jean-Michel Adam (Adam, 2011b). Este postulado se da dentro de lo que se conoce como lingüística textual, la cual parte de la lingüística del discurso de Émile Benveniste (2006), hasta el modelo de secuencias y prototipos de Adam (2011). Por último, encontramos el concepto de modalización, como lo han trabajado Patrick Charaudeau (1992) y Dominique Maingueneau (Charaudeau y Maingueneau, 2002), dentro de una teoría más general del análisis del discurso, pero que no se encuentra desligada de la lingüística textual, como propone Adam en la tercera edición de *La linguistique textuelle: introduction à l'analyse textuelle des discours* (Adam, 2011b). Si bien estos tres elementos –el género discursivo, los planos (o niveles) textuales y la modalización– hacen parte de la lingüística y del análisis del discurso, nos interesa destacar que son tradiciones que se han preocupado por los problemas y la naturaleza de los textos desde su relación con los géneros discursivos y, en esa medida, con las condiciones sociales de la interlocución textual, y no tanto por el análisis sintáctico, conversacional o exclusivamente pragmático.

3.1. Sobre el género discursivo

En cuanto al género discursivo, hay un enfoque que procura tratarlo como parte de situaciones comunicativas que conllevan varios niveles y módulos no jerárquicos. En el caso de Adam (2011b), se trata de pensar cómo el problema del género discursivo empezó a entrar en una tradición lingüística que había mantenido la frase como unidad máxima de análisis. La constatación de la diferencia de índole entre la frase y el texto hace necesario el planteamiento de una nueva teoría y de otros métodos para estudiar el funcionamiento textual y, con relación a él, la constitución de los géneros. La pregunta central de esta rama reside en cómo trabajar el género dentro de la lingüística textual y analizar qué papel puede tener. Los principales aportes de esta perspectiva teórica para nuestro estudio residen en dos puntos concretos: en primer lugar, pensar

⁵ Al respecto de estos dos términos, seguimos a Pinceman: “La textometría es una nueva forma de nombrar la lexicometría, para dar cuenta de mejor manera de las evoluciones recientes de esta disciplina (sobre todo en el tratamiento de corpus etiquetados) y evitar que se entienda que estos estudios se limitan a un estudio léxico”. (Pincemin 2011:54, nuestra traducción).

el género discursivo en términos de *efectos de genericidad* y, en segundo lugar, la comprensión de que los distintos planos de organización textual están directamente determinados tanto por la lengua como por la genericidad. Según Adam y Heidmann,

desde que hay un texto –es decir, el reconocimiento del hecho de que una serie de enunciados forman un todo comunicativo–, hay un efecto de *genericidad*, es decir, la inscripción de esa serie de enunciados en una clase de discurso. La *genericidad* es una necesidad socio-cognitiva que vincula cualquier texto al interdiscurso de una formación social. Un texto no pertenece, en sí, a un género, sino que es puesto en relación, tanto en su producción como en su recepción e interpretación, con uno o con varios géneros (Adam y Heidmann, 2004: 62, nuestra traducción).

La genericidad así entendida hace posible analizar el género más allá de un repertorio de características más o menos normativo. Por el contrario, nos permite asumir la complejidad de los procesos de producción y de recepción de los textos y de los géneros discursivos en términos de genericidad, así como de efectos de genericidad. En otras palabras, como hablantes e integrantes de ciertas comunidades discursivas, reconocemos un efecto de *textualidad* cuando identificamos los enunciados de un texto como un todo, e identificamos la *genericidad* cuando, al producir o al recibir textos, los vinculamos a formaciones sociodiscursivas que consideramos como géneros. Los artículos de investigación que conforman el corpus de este estudio serán trabajados desde esta perspectiva que procura vincular dos características: su dimensión de necesidad cognitiva –que enmarca la producción y la recepción de textos–, por un lado, y su carácter social y discursivo dentro de comunidades dadas, por otro.

En un cruce entre un paradigma cognitivo y una rama social de estudios del discurso, Adam y Heidmann nos orientan hacia un concepto de género que contempla la situación comunicativa en la producción y en la recepción de textos de una manera dinámica y no normativa ni enmarcada en la enseñanza de las lenguas.

3.2. Sobre el nivel de la enunciación

Dentro de esta tradición, y por medio de una crítica explícita a las tipologías escolares de textos, Adam plantea un esquema de niveles de organización de la textualidad. Este será el segundo elemento conceptual que emplearemos: los niveles textuales. En su propuesta, Adam parte de la inscripción de los distintos niveles textuales en una dimensión más amplia de los géneros discursivos: estos últimos serían los únicos que constituirían algún tipo de categorización superior con respecto al texto. Como lo afirma el autor al cerrar el prefacio de la edición de 2011:

Reafirmo sin ambigüedad que aparte de lo que pude escribir en los años ochenta, para mí el concepto de tipos de textos es, como *la gramática del texto*, más un obstáculo metodológico que una herramienta heurística. Las tipologías fundadas sobre la genericidad son mucho más pertinentes, porque es en gran medida el género discursivo –dependiendo él mismo de los marcos de interacción sociodiscursivos– el que determina la composición de los textos. Los géneros discursivos son los únicos hechos de alto nivel susceptibles de ser categorizados. (Adam, 2011b: Capítulo Avant-propos, nuestra traducción)

En el marco teórico para la propuesta de las secuencias textuales prototípicas que expone y elabora Adam –a partir de la lingüística textual y de los postulados de Bajtín–, el lingüista francés construye un modelo de las interacciones entre los niveles o fases textuales, y entre estos y las configuraciones genéricas y discursivas más generales. Reproducimos aquí la versión más reciente de este esquema, puesto que emplearemos en particular uno de los niveles propuestos por Adam:

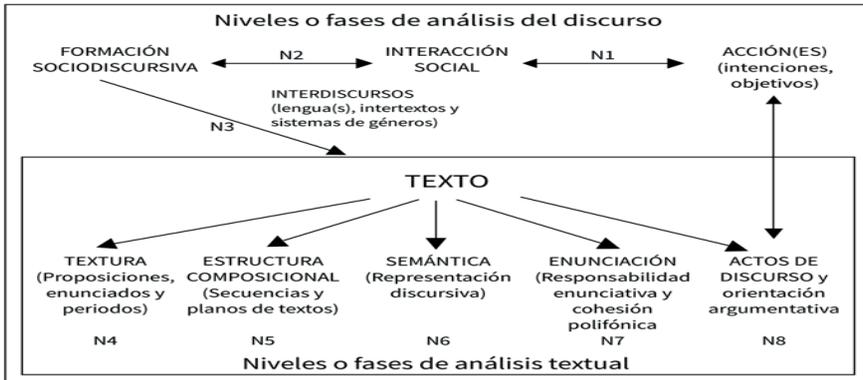


Figura 1. *Niveles del análisis de discurso y niveles de análisis textual de Jean-Michel Adam (Adam 2011b), nuestra traducción.*

En total, la propuesta cuenta con ocho niveles o fases que funcionan de manera modular y no jerárquica. Los primeros tres corresponden a la dimensión discursiva (extratextual e intertextual) en la que se inscribe cualquier texto. Los siguientes cinco corresponden a niveles textuales de análisis. Para este artículo es relevante en particular el séptimo nivel (N7), el de la enunciaci3n, el cual involucra la responsabilidad frente a los enunciados y la cohesi3n polif3nica (es decir, la articulaci3n con otras voces citadas o vinculadas al texto). Puesto que todos los art3culos del corpus contienen estos niveles textuales, y dado que las formas de an3lisis de cada nivel pueden divergir, precisamos que la primera entrada al corpus a partir de las marcas de modalizaci3n implica el estudio de un nivel particular de an3lisis: el de la enunciaci3n (N7). En adelante, nos centraremos en el an3lisis del corpus a partir de ese nivel enunciativo (N7) en el que es posible observar la responsabilidad sobre los enunciados y, en especial, la modalizaci3n que quien enuncia hace sobre sus enunciados, su interlocutor y s3 mismo.

3.3. Sobre la modalizaci3n

A partir del nivel de la enunciaci3n (N7) de Adam (2011a), el equipo pas3 a indagar, dentro del corpus ya establecido, c3mo operaban las marcas de modalizaci3n frecuentes y qu3 efectos produc3an, as3 como cu3les de ellas apuntaban a una implicaci3n del sujeto de la enunciaci3n dentro de los art3culos. Por modalizaci3n entendemos un elemento que, siguiendo a Patrick Charaudeau, “no solo constituye una parte del

fenómeno de la enunciación, sino que es su pivote en la medida en que permite explicitar las posiciones del sujeto hablante con respecto a su interlocutor, a sí mismo y a lo que dice” (Charaudeau, 1992: 572, nuestra traducción). Puesto que las marcas de modalización permiten rastrear la posición y la actitud de quien enuncia, estas fueron tomadas como una primera entrada al corpus (en especial a partir de segmentos repetidos, como se detallará en la metodología) y al análisis de las formas de implicación que los investigadores pueden o se permiten tener frente a sus artículos de investigación y a su propio rol dentro de los textos.

El grado o nivel de explicitación de posiciones es planteado aquí como una primera forma de apuntar si estamos ante la construcción de efectos de despersonalización –o de distanciamiento u ocultamiento– de las actitudes de quienes enuncian, los cuales, más adelante, ayudarían a esbozar una concepción de investigación determinada. En este artículo, planteamos estos fenómenos en términos de efectos que producen los textos a partir de su enunciación –y particularmente de su modalización–, y no en función de la agencia real de los investigadores en sus estudios, o de su subjetividad u objetividad a la hora de investigar. Es decir, partimos de la distinción entre el ámbito de la enunciación y el de las prácticas investigativas: acá estudiaremos el primero, con especial énfasis en las marcas de modalización.

No sobra recordar en este punto que, siguiendo a Charaudeau, entendemos por marcas alocutivas aquellas que implican al interlocutor; por elocutivas aquellas en las que el locutor revela su posición sobre lo que dice; y por delocutivas aquellas en las que el locutor se deja imponer enunciados como si no fuese el responsable de su enunciación, es decir, son marcas en las que el locutor y el interlocutor parecen estar ausentes (Charaudeau, 1992, p. 576). En todas está presente el sujeto, solo que esa presencia se da en modalidades diferentes.

4. METODOLOGÍA

Puesto que el equipo de trabajo buscaba construir un corpus en español y de artículos publicados en Latinoamérica, se decidió emplear el índice citacional SciELO (Scientific Electronic Library Online) que visibiliza la producción intelectual de la región. Luego, se procedió a construir una ecuación de búsqueda para explorar los primeros resultados de artículos relacionados con el lenguaje en las ciencias sociales y en las humanidades⁶. Posteriormente, se realizó una depuración de los textos a partir de los resúmenes para garantizar que el corpus solo tuviese artículos pertinentes con respecto al tema. El resultado es un corpus de 1260 artículos publicados entre 2002 y

⁶ La ecuación de búsqueda fue: (Tema= Lenguaje AND Idioma= Spanish) y se aplicaron los filtros por categoría cognitiva de la revista de la siguiente manera: (humanities multidisciplinary OR linguistics OR philosophy OR language linguistics OR social sciences interdisciplinary OR cultural studies OR women s studies OR history of social sciences OR communication OR sociology OR anthropology OR literature OR history philosophy of science OR literary theory criticism). La fecha de esta consulta fue el 4 de septiembre de 2018, y el 15 de febrero de 2019 se actualizó la información.

2018. Después de la constitución de una base de datos inicial, se procedió a descargar los 1260 artículos y a pegarlos en un solo archivo .TXT que pudiera trabajarse en dos programas informáticos de lexicometría para análisis computarizado de lengua natural: Iramuteq y Lexico 5. La razón para elegir estos programas es que son de acceso libre y cuentan con manuales y tutoriales gratuitos en línea. Además de eso, Iramuteq permite estadísticas generales de las frecuencias de palabras, mientras que Lexico 5 permite el análisis de segmentos repetidos y una visualización de las coocurrencias más sencilla. Es importante aclarar que, para el presente estudio, el equipo privilegió un enfoque heurístico de corpus (*corpus-driven approach*) en el cual el análisis de los textos se lleva a cabo sin modelos teóricos ni hipótesis preestablecidas, pero sí con los criterios de selección de los artículos mencionados anteriormente.

Ahora bien, siguiendo el enfoque heurístico, el analista, guiado por el mismo corpus, elabora hipótesis interpretativas relacionadas con hechos lingüísticos allí encontrados (Rastier, 2011). En consonancia con esta orientación, durante varias sesiones de trabajo en 2019, el equipo empezó a observar las frecuencias de palabras más altas en Iramuteq y a compararlas con los segmentos repetidos más frecuentes en Lexico 5 (sobre todo en segmentos de seis, cinco y cuatro palabras). El uso de estos dos programas informáticos se debe a que son de acceso libre y cuentan con manuales de libre circulación. Cada sesión implicó discusiones acerca de los eventuales observables en el corpus que darían cuenta de dos aspectos: en primer lugar, de diferentes formas de modalización (especialmente en marcas delocutivas y elocutivas); en segundo lugar, de cómo las marcas de modalización revelan cómo operan las convenciones del género discursivo que conocemos como artículo de investigación en español.

Si bien la modalización es una categoría conceptual y no formal, y en esta medida, es implícita en el discurso, hay algunas marcas formales en los textos que sí pueden expresar lo alocutivo, elocutivo y delocutivo de manera explícita. En otras palabras, las marcas de modalización pueden ser analizadas en la superficie textual del corpus, a pesar de ser sincréticas (pueden funcionar para distintas intenciones de comunicación), polimorfos (pertenecen a sistemas formales diferentes) y no monosémicas (Charaudeau, 1992, p. 573). El análisis de los segmentos repetidos y del cotexto de estos hizo posible una primera observación de marcas delocutivas explícitas (y de algunas elocutivas) y su funcionamiento dentro del corpus.

A partir de los segmentos repetidos en Lexico 5, el equipo empezó a identificar, sobre todo en segmentos entre cuatro y seis palabras, cuáles apuntaban a un fenómeno de modalización, bien sea con relación al sujeto de la enunciación (elocutivo) o al enunciado (delocutivo). Los segmentos repetidos de más de seis palabras tendían a ser títulos de textos o nombres de proyectos y no tanto locuciones o expresiones de lengua común. Por esta razón, a partir de segmentos repetidos de seis palabras se inició el análisis de manera descendente (es decir, hacia cuatro palabras)⁷. Luego, al observar que la mayoría de segmentos repetidos entre seis y cuatro palabras apuntaba

⁷ Dado el tamaño del corpus, para este artículo solo presentamos el análisis de segmentos repetidos de seis, cinco y cuatro palabras.

a marcas delocutivas, se pasó a una segunda fase del análisis. En esta se establecieron una serie de categorías para agrupar los segmentos y describir así las distintas marcas delocutivas encontradas (y también las pocas elocutivas observadas).

De este modo, la primera aproximación a los segmentos repetidos se hizo a partir de la clasificación de Charaudeau (marcas alocutivas, delocutivas y elocutivas). Al encontrar la predominancia de las marcas delocutivas, en la segunda aproximación se optó por generar categorías propias, pero inspiradas en Charaudeau (1992) que describieran la función de cada conjunto de marcas. Luego de tener esa clasificación, un tercer momento consistió en tomar algunas de las palabras clave de los segmentos repetidos y contrastar su frecuencia con la información de Iramuteq. Esto permitía saber qué tan significativa era la aparición de una palabra determinada dentro de un segmento repetido con respecto a la frecuencia de aparición de la palabra sola (lematizada, es decir, con todas sus variantes) en todo el corpus. Para ilustrar esta explicación, podemos recurrir a un ejemplo concreto: el segmento repetido “No cabe duda de que” tiene una frecuencia de 60 eventos en Lexico 5, mientras que la palabra “duda” tiene una frecuencia de 1606 eventos (sin lematizar) y de 2043 en todo el corpus (lematizada, incluso como verbo también), según el tratamiento de Iramuteq. Estos tres momentos constituyeron las fases de análisis de las marcas de modalización y se representan gráficamente en la Figura 1. A continuación, presentaremos los hallazgos de esta primera aproximación a las marcas de modalización en el corpus.



Figura 2. Fases metodológicas

5. RESULTADOS

Los hallazgos de este estudio tienen que ver con dos aspectos: en primer lugar, con la dominancia de marcas delocutivas en los segmentos repetidos (analizados en Lexico 5); en segundo lugar, con la construcción de categorías que agrupan esas marcas delocutivas encontradas y con la alta frecuencia de algunas de ellas.

5.1. Se investiga: la dominancia de lo delocutivo

En cuanto al primer aspecto, es decir, a la dominancia de marcas delocutivas encontramos lo siguiente. Por una parte, las marcas alocutivas están ausentes de los segmentos repetidos analizados. Esto podría indicar que las marcas con una

función alocutiva parecen ser poco características del género discursivo del artículo de investigación. Esto confirma lo que algunos estudios han trabajado con corpus entre seis y cincuenta veces más pequeños (Fuentes Cortés 2012; Sabaj Meruane y otros, 2011). Lo anterior puede deberse a que hay pretensión de objetividad, pero sobre todo a que este género discursivo se conciba como poco dialógico. Dado que la comunicación y la interacción entre los autores y los lectores está mediada por largos procesos e, incluso, por tiempos muy amplios que superan a veces la vida de los autores, se trata de un género que deja pocas huellas de la interacción con sus lectores. A diferencia de géneros como las reseñas o las cartas, las marcas alocutivas parecen ser inexistentes (al menos en estos segmentos repetidos) y esto ayudaría a construir el efecto de despersonalización de un texto destinado a públicos amplios, periodos de recepción indeterminados y una lectura mediada por una gran distancia entre quien escribe y quien lee. Esto no significa que no haya posturas de autor: las hay y estas se presentan predominantemente de manera despersonalizada. Si la escritura tiene esta capacidad de objetivación, el artículo de investigación en español parece reforzarla por medio de esta inexistencia del rastro de los interlocutores directos. La distancia puede ser tal que no parece haber presencia de ningún interlocutor en locuciones o expresiones verbales frecuentes en este género discursivo, al menos en segmentos entre cuatro y seis palabras.

Por otra parte, en los segmentos analizados solo hay marcas elocutivas como las siguientes: “Nos damos cuenta de que”, “Desde mi punto de vista”, “Desde nuestro punto de vista” y “En mi opinión” con frecuencias entre 10 y 13 en todo el corpus. Estas cuatro expresiones contrastan con las 72 que responden a marcas delocutivas como, por ejemplo, “No hay que olvidar que” (20 eventos), “No cabe duda de que” (60 eventos), “A diferencia de lo que” (75 eventos), “A pesar de que (los, las, la, en, el, no, se)” (140 eventos) o “En la medida en que (los, las, la, lo, no, se, el, esta)” (386 eventos). Como se observa, los segmentos repetidos elocutivos son significativamente menos frecuentes con respecto a los delocutivos. Esto apunta a que las marcas delocutivas, al menos en segmentos repetidos, parecen ser dominantes en el corpus. Esta predominancia puede explicarse por la búsqueda de un efecto de despersonalización en este género discursivo. Si bien se trata de un género en el que se argumenta y se debate sobre posturas, autores y teorías, parece ser dominante el hecho de que esto ocurra empleando principalmente fórmulas y locuciones en las que el locutor parece no asumir la responsabilidad directa de los contrastes o de las afirmaciones y sus matices. Esto no significa que la enunciación responda a la objetividad o subjetividad buscada en las investigaciones como tales ni que los autores no estén presentes como sujetos, sino que, en el plano de la enunciación (N7) de Adam (2011a), es convencional y frecuente proyectar una imagen despersonalizada en el discurso científico. Parece ser una convención recurrente acudir a locuciones y expresiones fijas en ciertas circunstancias enunciativas que explicaremos por medio de la propuesta de una serie de categorías. Estas permitirán agrupar semánticamente los segmentos repetidos encontrados y argumentar en qué medida hablamos de un efecto de despersonalización.

5.2. Categorías propuestas

Lo anterior nos lleva al segundo aspecto de los resultados: el de las categorías propuestas. A raíz del análisis de los segmentos repetidos, se propusieron categorías discursivas. Estas se detallan en la siguiente tabla junto con las frecuencias respectivas. Cada categoría contiene varios segmentos repetidos que comparten un campo semántico más amplio. Por ejemplo, la categoría de “posibilidad” incluye los siguientes segmentos repetidos: “A la posibilidad de que”, “Por lo que se puede”, “La posibilidad de acceder a” y “La posibilidad de que”. La frecuencia de estos segmentos es de 112. La tercera columna de la tabla indica algunos ejemplos y la cuarta cuáles categorías fueron interpretadas como predominantemente delocutivas o elocutivas. Algunas categorías fueron tomadas de Charaudeau (Charaudeau y Maingueneau, 2002) y otras fueron propuestas a partir del análisis del corpus mismo, como lo sugiere Tognini-Bonelli (2001).

Marcadores	Frecuencias	Ejemplos de segmentos repetidos	Modalización
Generalización	884	“En la mayoría de”	Delocutiva
Condición	504	“En la medida en que (los, las, la, lo, no, se, el, esta)”	
Oposición	368	“A pesar de que”	
Posibilidad	112	“A la posibilidad de que”	
Deóntica	96	“Por lo que es necesario”, “Debe tenerse en cuenta que”	
Certeza	81	“Lo cierto es que”, “No cabe duda de que”	
Ilustración	60	“Como se puede ver en”, “Como se puede apreciar en”	
Énfasis	58	“Lo que es más importante”, “Por la sencilla razón”	
Relatividad	50	“En mayor o menor grado”	
Ampliación	40	“Van más allá”, “En un sentido más amplio”	
Evidencia	26	“Nos damos cuenta de que”	
Apreciación cuantitativa	29	“Con un alto grado de”	
Complejización	23	“Más allá de la mera”	
Opinión	22	“Desde mi punto de vista”	Elocutiva
Totalizante	20	“De la totalidad de los”	Delocutiva
Probabilidad	18	“Lo más probable es que”	
Concesión	13	“Si bien es cierto que”	
Inexistencia	10	“En ninguno de”	

Tabla 1. Categorías que agrupan los segmentos repetidos

A continuación, también exponemos una gráfica que permite apreciar el contraste de las frecuencias entre las distintas categorías. Las frecuencias de las primeras cuatro categorías están entre 112 (posibilidad) y 884 (generalización) en todo el corpus, lo cual nos hace pensar que se trata de marcas de modalización dominantes. Por ello, nos detendremos en su análisis.

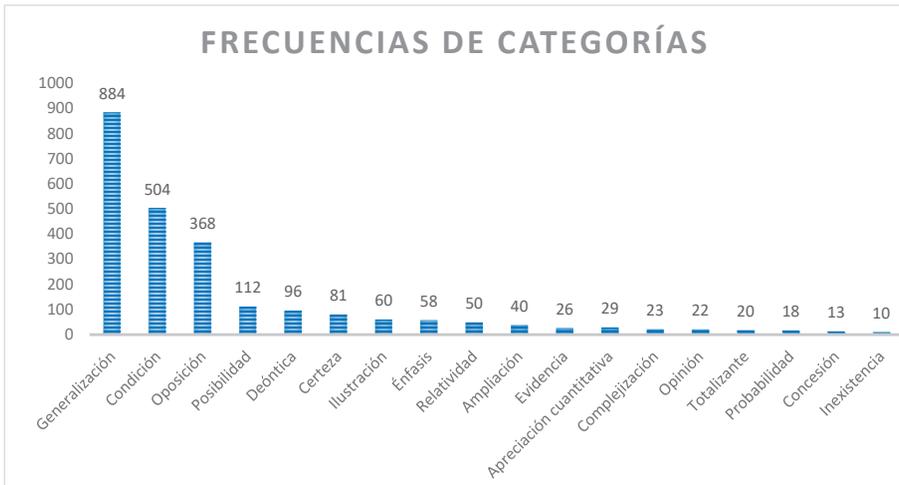


Figura 3. *Frecuencias de las categorías propuestas*

Como se observa en la Figura 3, los segmentos repetidos que emergen con mayor frecuencia en el corpus tienen que ver con cuatro categorías: la generalización, la condición, la oposición y la posibilidad. La generalización parece ser una categoría que opera con frecuencias bastante altas en los segmentos repetidos. Es importante destacar aquí que todos los segmentos corresponden a marcas delocutivas, es decir, se generaliza sin explicitar quién es el responsable de esa generalización. También cabe señalar que no se trata de una generalización totalizante, sino indeterminada cuantitativamente: las frecuencias de la categoría de “Totalización” (10 para “A la totalidad de la” y 10 para “De la totalidad de los”) son significativamente más bajas en todo el corpus. Esto puede significar que en el artículo de investigación opera la convención de generalizar, pero de manera no totalizante, sino indefinida. Lo anterior es coherente con la reiteración de la palabra “mayoría” y “mayor” en los segmentos repetidos (Figura 4).

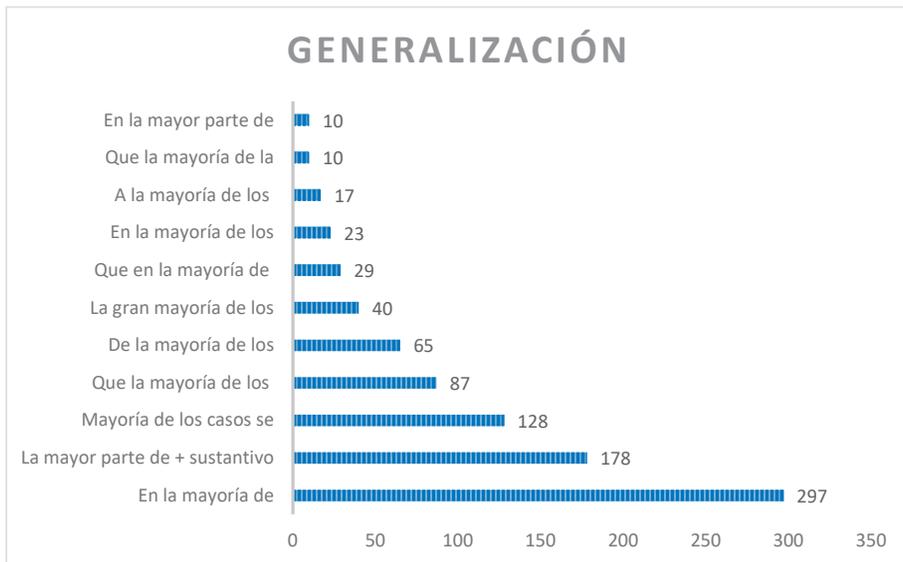


Figura 4. Frecuencias de segmentos repetidos de “generalización”

Ahora bien, llama la atención que la frecuencia más alta la tengan segmentos repetidos que agrupamos en el campo semántico de la generalización. El artículo de investigación, como parte de los géneros académicos e investigativos, es concebido generalmente como un género que busca convencer a una comunidad discursiva especializada (Björk y otros, 2003; Charaudeau y Maingueneau, 2002). En este sentido, es llamativo que uno de los efectos retóricos de estos textos se dé a través del uso habitual de generalizaciones. Esto apunta a que es frecuente (y quizá más eficiente en términos de un propósito comunicativo de convencimiento) generalizar ciertos asuntos (resultados, datos, elementos de contexto) con el fin de darle relevancia y justificación a aquello que se dice. Es notoria la alta frecuencia (297) del segmento “En la mayoría de”, como una fórmula fija que tal vez permita manifestar esa generalización en distintos apartados de los textos.

La segunda categoría que presenta las frecuencias más altas es la “condición”. En ella encontramos tanto segmentos con los que inician oraciones condicionales (“Si se tiene en cuenta”) como el segmento con la frecuencia más alta “En la medida en que” (386 eventos). A diferencia del uso de un cuantificador indeterminado como “la mayoría”, el uso reiterado de este segmento podría apuntar a condicionar las afirmaciones y, en esta medida relativizarlas, restringirlas o matizarlas. Esta restricción se presenta como una marca delocutiva que limita y condiciona la proposición que acompaña, ya sea en términos de resultados expuestos, del contexto o el planteamiento del problema, la elaboración de conclusiones o la explicación de un marco teórico. La alta frecuencia de marcadores de condición en todo el corpus podría explicarse por el carácter argumentativo del artículo de investigación, puesto que los argumentos que se desarrollan no son absolutos y tenderían a evitar las simplificaciones. Tal vez

por esa razón, el uso del segmento “En la medida en que” sea una marca que permite condicionar y restringir el significado de ciertas ideas o argumentos, con el fin de ganar rigor argumentativo, incluyendo y anticipando posibles reparos o críticas a ideas determinadas o que podrían tomarse como tajantes. De nuevo, es llamativo que esta función argumentativa de la condición se exprese predominantemente por medio de marcas delocutivas en este corpus (Figura 5). La condición se expone en el conjunto del corpus como algo que el discurso mismo impone, no como una responsabilidad de quien escribe. Esto podría ser determinante en el efecto persuasivo del género “artículo”: precisamente porque parece despersonalizada la argumentación, puede funcionar como si se tratara más una muestra de evidencias que de una cuestión de argumentos. En otras palabras, el convencimiento se refuerza por aparentar no convencer, sino por imponerse en la enunciación como algo que se muestra.



Figura 5. Frecuencias de segmentos repetidos de “condición”

La tercera categoría con una frecuencia alta es la de “oposición”. Si bien estos segmentos repetidos son esperables en un texto que se considera predominantemente argumentativo, es importante señalar que su frecuencia (368 eventos) es inferior por más de la mitad con respecto a las marcas de generalización. Al menos en el análisis de estos segmentos repetidos, parece emplearse más la generalización que las marcas de oposición entre ideas –marcas que, por otro lado, tenderíamos a asociar más al carácter polémico o de disenso que implica todo texto argumentativo–. Si planteáramos una hipótesis inicial sobre los artículos como textos argumentativos, podría esperarse que estas marcas tuvieran frecuencias más altas. Su presencia en el corpus es bastante significativa, pero quizá el efecto argumentativo no se observa solamente en este tipo de marcas que oponen elementos, ideas, autores, teorías o argumentos. Es destacable que la locución “a pesar de” sea dominante, por encima de otras como “lo contrario de lo que” o “a diferencia de” (Figura 6). Esto quizá se deba al uso de un registro considerado más culto en el que la crítica y la oposición se enuncian de manera menos directa o explícita para los hablantes, lo que no ocurriría con palabras como “contrario” o “diferencia”, en las que la oposición resulta tal vez más nítida.

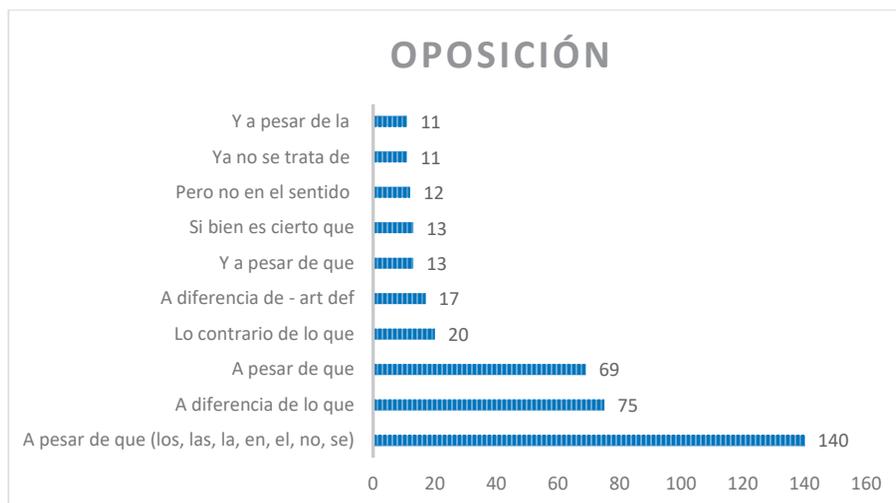


Figura 6. Frecuencias de segmentos repetidos de oposición

La cuarta categoría con frecuencias altas agrupa segmentos repetidos que tienen en común lo que denominamos “posibilidad”. La frecuencia total de esta categoría es de 112 e incluye una variedad menor de segmentos repetidos, puesto que son solo cuatro. Destacamos también aquí el segmento “La posibilidad de que” con una frecuencia de 70, puesto que puede apuntar un carácter hipotético o matizado de las formulaciones en este género discursivo. Para verificar este resultado en todo el corpus, y dado que las frecuencias no son tan altas, sería necesario un cruce de estos datos con otras marcas de modalización que se refieran a la expresión de la “posibilidad”, como el uso del modo subjuntivo, por ejemplo, a través de Iramuteq. En otro artículo, desarrollamos un análisis de las formas verbales (Lamprea Abril y otros, en prensa).

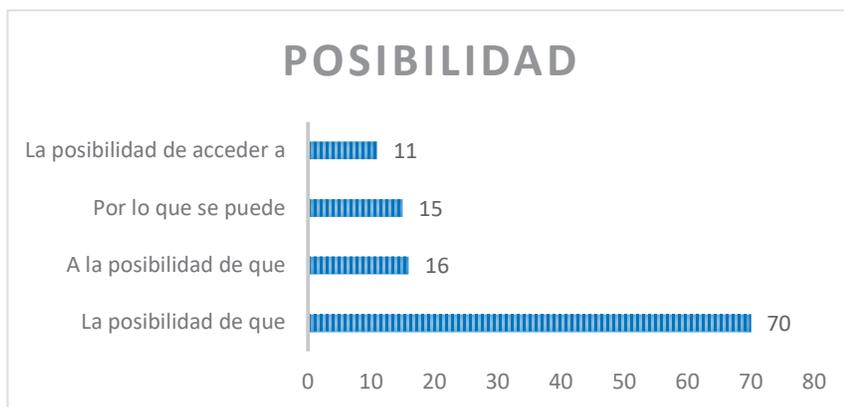


Figura 7. Frecuencias de segmentos repetidos de “posibilidad”

6. CONCLUSIONES

Para concluir, destacaremos elementos de estos resultados preliminares vinculados con tres aspectos: el funcionamiento de la modalización en los segmentos repetidos y su relación con el género discursivo “artículo”; la posible relación de estos hallazgos con una concepción de investigación; y una observación metodológica. En cuanto a la modalización y su relación con este género, cabe insistir en la dominancia del carácter delocutivo de estas marcas de modalización que organizamos en categorías. Es significativo que la generalización, la condición, la oposición y la posibilidad, por ejemplo, agrupen segmentos delocutivos: esto puede implicar que, a pesar de que los artículos pueden expresar subjetividad, como señala García Negroni (2008), los segmentos repetidos más frecuentemente utilizados en este corpus revelan la presunta ausencia de quien enuncia. En otras palabras, una de las convenciones de este género tiene relación con el uso de fórmulas y locuciones predominantemente delocutivas, es decir, con expresiones más o menos fosilizadas que expresan una actitud frente al contenido enunciado, pero sin que la responsabilidad de la enunciación sea visible y explícita. Lo anterior no implica que haya neutralidad u objetividad del discurso, sino que la actitud de quien enuncia parece orientarse más hacia lo que se dice que hacia quien lo dice, como en “La posibilidad de que” o “A diferencia de lo que”. En las locuciones más fijas del género “artículo” parece ser más relevante la actitud frente a lo enunciado (ya sea deóntica, de certeza, de generalización u otra), que la actitud frente a quien enuncia. Si bien en apariencia puede ser un rasgo evidente para la lingüística aplicada, para los enfoques normativos o de escritura académica, nos parece que, dado el tamaño del corpus, es un resultado significativo que habría que seguir estudiando tanto en sus formas actuales como en su historia en los países hispanohablantes. ¿Qué nos dice de la enunciación frente a la investigación y al conocimiento? ¿Qué tanta consciencia del papel del sujeto de la enunciación expresa esta actitud preponderante frente a lo enunciado? ¿Funciona de maneras similares en distintas lenguas? ¿Qué tanto el artículo de ciencias naturales (y su publicación en lengua inglesa) como hegemónico ha permeado el género en español y la dominancia de estas marcas?

Lo anterior se vincula con nuestra pregunta inicial, puesto que esta implica la indagación por las marcas lingüísticas y textuales de la modalización. Con respecto a García Negroni (2008) en particular, es importante afirmar que las marcas de modalización delocutiva encontradas confirman una “despersonalización” en el género “artículo”. Este efecto podría corroborarse con las frecuencias de los pronombres personales lematizados en Iramuteq, que arrojan una frecuencia de 40 722 para el pronombre “él”, 11 463 para “nosotros” y 10 434 para “yo”. No obstante, esta hipótesis tendría que confirmarse con un futuro análisis de esas frecuencias y su contraste con las de los verbos, dada la frecuente elisión de los sujetos gramaticales en español.

También es relevante insistir en que el efecto de despersonalización ocurre sobre todo en segmentos que representan expresiones o locuciones que están ya fosilizadas como “en la medida en que” o “a pesar de que”. En el caso de las frecuencias de palabras aisladas, las formas de modalización podrían mostrar otros resultados, como

la frecuencia de “podemos” (3230), “encontramos” (1162) o “creemos” (630), en los cuales podría haber una función elocutiva. Esto significa que los segmentos repetidos dan cuenta de locuciones y expresiones fosilizadas predominantemente delocutivas, pero que quizá las frecuencias absolutas de otros elementos lingüísticos apunten a marcas elocutivas características del género. Estas últimas se limitaron a cuatro segmentos repetidos (clasificados como “opinión” y uno como de “evidencia” – “Nos damos cuenta de que”–), pero quizá su expresión se dé más en las desinencias verbales o en otros giros textuales que habrá que analizar más adelante en este mismo corpus.

Quizá las marcas elocutivas y el efecto de un discurso personal sean dominantes en momentos determinados de los textos y respondan a expresiones menos fosilizadas y a marcas de estilo propio que, por esa misma razón, se observarían en las frecuencias más bajas. En cualquier caso, la dimensión cognitiva y social (Adam y Heidmann, 2004) de este género parece determinar el uso frecuente de locuciones y expresiones delocutivas, es decir, se trata de un género que quizá haga énfasis en sus contenidos por medio de estas marcas relativamente fijas. Lo interesante es que ese énfasis delocutivo no impide que haya un posicionamiento de los autores: por el contrario, las locuciones deónticas, de certeza, oposición o de énfasis, sí implican que cada autor se sitúe frente a su tradición. Sin embargo, ese posicionamiento no se hace visible en términos de la expresión personal, sino como una actitud frente a los enunciados mismos. Se trata de mostrar las actitudes y posiciones frente a los enunciados y frente al discurso mismo de las disciplinas, no frente a los sujetos de la enunciación.

Ahora bien, ¿qué nos dicen estos resultados frente a posibles concepciones de investigación? A pesar de que en este artículo el corpus fue tratado sin divisiones disciplinares, es posible afirmar que se encuentra una concepción de investigación que parece darle más relevancia al efecto de despersonalización del discurso que a la visibilidad de quien enuncia lo investigado. Esto no significa que predomine la argumentación o la exposición, sino que, en cualquiera de los dos casos, el género “artículo” parece operar a partir de una convención que les da un lugar importante a las marcas delocutivas, por medio de locuciones fijas. En esta medida, la investigación parece concebirse aquí desde una postura que privilegia la impersonalidad del saber, al menos en la manera en que se formula (lo cual no quiere decir que los autores no tomen posturas). Es más frecuente que se discutan los enunciados –y también la dimensión discursiva que involucra sus condiciones de producción– a que se hagan visibles las actitudes de los autores sobre sí mismos. Esto podría ayudar a describir mejor el tipo de argumentación que caracteriza este género discursivo: se trata de una argumentación que se teje a través de efectos de despersonalización por medio del uso de locuciones más o menos estables. Así, desde las categorías propuestas, si concebimos las marcas de modalización en relación con prácticas sociales, estas podrían señalar posicionamientos frente a lo que se piensa que es investigar. En ellas, investigar se asocia con la predominancia de, por ejemplo, la generalización de lo enunciado y la oposición frente a otras posturas para matizar lo que se dice, desde una postura que no muestra su agencia. Esta sería una posible lectura de cómo este análisis concreto de las marcas textuales de modalización puede vincularse con prácticas sociodiscursivas de lo que se valora como investigación.

Para terminar, quisiéramos apuntar una conclusión metodológica. El hecho de formular hipótesis a partir del corpus mismo implica cierta prudencia en el tipo de afirmación que se elabora. En este sentido, creemos fundamental que el análisis pueda contrastar frecuencias absolutas de palabras aisladas (en Iramuteq, por ejemplo) con la frecuencia de los segmentos repetidos (en Lexico 5): los datos en un caso y otro pueden implicar aseveraciones distintas. Por ejemplo, la frecuencia del segmento “en ninguno de (los, las)” fue solo de 10, mientras que la palabra “ninguno” lematizada aparece 1806 veces y “ningún” 1299 veces. Sin el contraste, el primer resultado llevaría a pensar que el corpus es matizado o poco taxativo con respecto a la inexistencia; el segundo, al contrario, indicaría que la afirmación de la inexistencia es frecuente y determinada. En este sentido, para quienes exploren grandes corpus para análisis similares al nuestro, es pertinente el cruce de este tipo de información. En cualquier caso, consideramos que el enfoque de la lingüística de corpus puede ser relevante para los estudios en español, dado que permiten contrastar diversas hipótesis acerca del funcionamiento de textos reales y de la categoría sociocognitiva del género, más allá de la lectura detallada de muestras representativas de textos.

7. REFERENCIAS

- ADAM, JEAN-MICHEL. 1997. “Genres, textes, discours: pour une reconception linguistique du concept de genre”. *Revue belge de philologie et d’histoire* 75(3):665–81.
- ADAM, JEAN-MICHEL. 2011a. *La linguistique textuelle. Introduction à l’analyse textuelle du discours*. 3e ed. Paris: Armand Colin.
- ADAM, JEAN-MICHEL. 2011b. *Les textes, types et prototypes : séquences descriptives, narratives, argumentatives, explicatives, dialogales et genres de l’injonction-instruction*. 3e éd. Paris: A. Colin.
- ADAM, JEAN-MICHEL, y UTE HEIDMANN. 2004. “Des genres à la généricité. L’exemple des contes (Perrault et les Grimm)”. *Langages* 38(153):62–72.
- BAJTÍN, MÍJAIL. 1999. “El problema de los géneros discursivos”. Pp. 248–93 en *Estética de la creación verbal*. México, D. F.: Siglo XXI Editores.
- BAZERMAN, CHARLES, ADAIR BONINI, y DÉBORA FIGUEREDO, eds. 2009. *Genre in a Changing World*. Fort Collins and West Lafayette.
- BENVENISTE, ÉMILE. 2006. *Problèmes de linguistique générale, 1*. Paris: Éditions Gallimard.
- BJÖRK, LENNART, GERD BRÄUER, LOTTE RIENECKER, y PETER STRAY JÖRGENSEN. 2003. *Teaching Academic Writing in European Higher Education*. Dordrecht: Kluwer Academic Publishers.
- BOLÍVAR, ADRIANA. 2004. “Análisis crítico del discurso de los académicos”. *Signos* 37(55):7–18.
- BOLÍVAR, ADRIANA. 2005. “Tradiciones discursivas y construcción del conocimiento en las humanidades”. *Signo y Seña* 14:67–91.
- CHARAUDEAU, PATRICK. 1992. *Grammaire du sens et de l’expression*. Paris: Hachette.
- CHARAUDEAU, PATRICK, y DOMINIQUE MAINGUENEAU. 2002. *Dictionnaire d’analyse du discours*. Paris: Éditions du Seuil.

- FUENTES CORTÉS, MIGUEL. 2012. “El discurso científico de la historia: Análisis estructural y retórico de los artículos de investigación en historia”. *Boletín de Filología* XLVII(1):89–110.
- FUENTES CORTÉS, MIGUEL. 2013. “Las conclusiones de los artículos de investigación en historia”. *Literatura y Lingüística* 28:215–32.
- GARCÍA NEGRONI, MARÍA MARTA. 2005. “Argumentación y polifonía en el discurso científico-académico. A propósito de ciertos conectores especializados en la trasgresión argumentativa”. *Rasal* 1:11–24.
- GARCÍA NEGRONI, MARÍA MARTA. 2008. “Subjetividad y discurso científico-académico. Acerca de algunas manifestaciones de la subjetividad en el artículo de investigación en español”. *Signos* 41(66):9–31.
- GARCÍA NEGRONI, MARÍA MARTA, ANDREA ESTRADA, y SILVIA RAMÍREZ GELBES. 2006. “Polifonía y objetividad. Las no coincidencias del decir en el abstract científico”. Pp. 378–89 en *Hispanismo: Discursos culturales, identidad y memoria*, editado por N. M. Flawiá de Fernández y S. P. Israilev. Tucumán: FFyLUNT.
- GARCÍA NEGRONI, MARÍA MARTA, y SILVIA RAMÍREZ GELBES. 2005. “Ethos discursivo y polémica sin enfrentamiento. Acerca del discurso académico en humanidades. En T. Rösing & C. Schons”. Pp. 114–37 en *Questões da escrita*, editado por T. Rösing y C. R. Schons. Passo Fundo: UPF.
- GARCÍA NEGRONI, MARÍA MARTA, y MARTA TORDESILLAS. 2001. *La enunciación en la lengua. De la deixis a la polifonía*. Madrid: Gredos.
- GONZÁLEZ ARIAS, CRISTIAN. 2011. “La formulación de los objetivos en artículos de investigación científica en cuatro disciplinas: historia, lingüística, literatura y biología”. *Linguagem em (Dis)curso, Tubarão* 11(2):401–29.
- LAMPREA ABRIL, NATHALIA, ANDREA TORRES PERDIGÓN, y DIANA MORENO RODRÍGUEZ. (en prensa). “¿Qué nos dicen los verbos acerca de lo que consideramos como investigación? Estudio de la modalización verbal en artículos de investigación relacionados con el lenguaje”. *Signos. Estudios de Lingüística*.
- MAYAFFRE, DAMON. 2005. “Rôle et place du corpus en linguistique. Réflexions introductives”. Pp. 5–17 en *Actes du colloque JETOU'2005*, editado por P. Vergély. Université de Toulouse-Le Mirail.
- MORENO, FRANCISCO, NORMA MARTHE, y LUIS ALBERTO REBOLLEDO. 2010. *Cómo escribir textos académicos según normas internacionales : APA, IEEE, MLA, Vancouver e Icontec*. Barranquilla (Colombia): Universidad del Norte.
- MOSCOVICI, SERGE. 1989. “Des représentations collectives aux représentations sociales : Éléments pour une histoire”. en *Les représentations sociales*, editado por D. Jodelet. Paris: Presses Universitaires de France.
- MURILLO, F. JAVIER, CYNTHIA MARTÍNEZ-GARRIDO, y GUILLERMINA BELAVI. 2017. “Sugerencias para escribir un buen artículo científico en educación”. *Revista Electrónica iberoamericana sobre calidad, eficacia y cambio en educación* 15(3):5–34.
- PINCEMIN, BÉNÉDICTE. 2011. “Analyse stylistique différentielle à base de marqueurs et textométrie.” Pp. 54–61 en *Vers une automatisation de l'analyse textuelle*. Vol. XX, editado por N. GARRIC y H. MAUREL-INDART. Texto ! Textes et Cultures, Institut Ferdinand de Saussure.
- RASTIER, FRANÇOIS. 2011. *La mesure et le grain. Sémantique de corpus*. Paris: Honoré Champion Éditeur.

- SABAJ MERUANE, OMAR, y DENISSE LANDEA BALIN. 2012. "Descripción de las formas de justificación de los objetivos en artículos de investigación en español de seis áreas científicas". *Onomázein* 23:315–44.
- SABAJ MERUANE, OMAR, PAULINA TORO TRENGOVE, y MIGUEL FUENTES CORTÉS. 2011. "Construcción de un modelo de movidas retóricas para el análisis de artículos de investigación en español". *Onomázein* 24(2):245–27.
- SABAJ, OMAR, MIGUEL FUENTES, y KEN MATSUDA. 2013. "Accidentes del verbo en las secciones de artículos de investigación". *Estudios filológicos* 52:129–42.
- SALAS VALDEBENITO, MILLARAY. 2015. "Una propuesta de taxonomía de marcadores metadiscursivos para el discurso académico-científico escrito en español". *Revista Signos. Estudios de Lingüística* 48(87):95–120.
- SANTESTEBAN-ECHARRIA, OLGA, y NURIA I. NÚÑEZ-MORALES. 2017. "Santesteban-Echarri, O., & Núñez-Morales, N. I. (2017). Cómo escribir un artículo científico por primera vez". *Psiquiatria Biologica* 24(1):3–9.
- SUÁREZ RELINQUE, CRISTIAN, GONZALO DEL MORAL ARROYO, y MANUEL TOMÁS FERNÁNDEZ GONZÁLEZ. 2013. "Consejos prácticos para escribir un artículo cualitativo publicable en Psicología". *Psychosocial Intervention* 22(1):71–79.
- SWALES, JOHN. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- SWALES, JOHN. 2004. *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- TOGNINI-BONELLI, ELENA. 2001. *Corpus linguistics at work*. Amsterdam: John Benajmins.