

Revista Española de Lingüística

Órgano de la Sociedad Española de Lingüística

RSEL

55/2

julio-diciembre 2025

REVISTA ESPAÑOLA DE LINGÜÍSTICA
(RSEL)

55/2

Edita
SEL

<https://doi.org/10.31810/RSEL.55.2>

REVISTA ESPAÑOLA DE LINGÜÍSTICA (RSEL)

ISSN: 0210-1874 • eISSN: 2254-8769

Depósito Legal: M-24.769-1971

DIRECTORA: Victoria Escandell-Vidal (Universidad Complutense de Madrid)

SECRETARIA: Marianela Fernández Trinidad (Universidad Complutense de Madrid)

RESPONSABLE DE RESEÑAS: Ventura Salazar (Universidad de Jaén)

CONSEJO EDITORIAL

Pilar Barbosa (Universidade do Minho)

Martin Becker (Universität zu Köln)

Virginia Bertolotti (Universidad de la República, Uruguay)

José A. Camacho (University of Illinois Chicago)

Ángeles Carrasco Gutiérrez (Universidad Castilla-La Mancha)

Concepción Company Company (Universidad Nacional Autónoma de México)

Bert Cornillie (KU Leuven)

Sonia Cyrino (Universidade Estadual de Campinas)

Eladio Duque Gómez (Universidad Complutense de Madrid)

Ricardo Etxepare (Centre National de la Recherche Scientifique)

Antonio Fábregas (University of Trondheim)

Aitor García Moreno (Consejo Superior de Investigaciones Científicas)

Eveling Garzón Fontalvo (Universidad de Salamanca)

Chiara Gianollo (Università di Bologna)

Kees Hengeveld (Universiteit van Amsterdam)

M.ª Dolores Jiménez López (Universidad Rovira i Virgili)

Johannes Kabatek (Universität Zürich)

Brenda Laca (Universidad de la República, Uruguay)

Jingsheng Lu (Shanghai International Studies University)

Rafael Marín Gálvez (Centre National de la Recherche Scientifique)

Pedro Martín Butragueño (El Colegio de México)

José Ignacio Hualde (University of Illinois Urbana-Champaign)

Victoria Marrero Aguiar (Universidad Nacional de Educación a Distancia)

Emilia Ruiz Yamuza (Universidad de Sevilla)

Begoña Sanromán Vilas (Helsingin Yliopisto / University of Helsinki)

Michelle Sheehan (Newcastle University)

Rui Sousa-Silva (Universidade do Porto)

Esperanza Torrego Salcedo (Universidad Autónoma de Madrid)

Luis Unceta Gómez (Universidad Autónoma de Madrid)

Axelle Vatrican (Université de Toulon).

COMITÉ DE HONOR

Juan Antonio Álvarez Pedrosa (Universidad Complutense de Madrid)
José Antonio Berenguer (Consejo Superior de Investigaciones Científicas)
Alberto Bernabé (Universidad Complutense de Madrid)
Margarita Cantarero (Sociedad Española de Lingüística)
Ramón Cerdá (Universidad de Barcelona)
Salvador Gutiérrez Ordóñez (Universidad de León y Real Academia Española)
Emma Martinell (Universidad de Barcelona)
José Carlos de Torres (Sociedad Española de Lingüística)

Los textos enviados para su publicación han de presentarse a través de la web: <http://revista.sel.edu.es/index.php/revista/about/submissions>. Deben ser originales e inéditos y ajustarse a las normas que aparecen en la página web de la *Revista Española de Lingüística*. Todos los artículos son sometidos al dictamen de al menos dos evaluadores (designados por el Consejo Editorial) que deben emitir informes de carácter confidencial.

Los derechos de publicación y difusión, bajo cualquier forma, son propiedad de la *RSEL*. Todo texto publicado en la revista obliga a sus autores a no cederlo a terceros, sin autorización previa de la revista. Todos los números de la revista se pueden encontrar en abierto en la web <<http://revista.sel.edu.es/>>.

SERVICIOS DE INFORMACIÓN: Los contenidos de la *RSEL* son recogidos sistemáticamente en *Bibliographie Linguistique/Linguistic Bibliography*, *Cindoc–Base de datos Sumarios Isoc*, *Dialnet*, *Doaj*, *Dulcinea*, *Circ*, *Latindex 2.0*, *Erih Plus*, *Dice*, *Cabells*.

REDACCIÓN: Sociedad Española de Lingüística, Centro de Ciencias Humanas y Sociales del CSIC, c/ Albasanz, 26-28, 28037 Madrid.

CORREO ELECTRÓNICO Y CORRESPONDENCIA: *RSEL.secretaria@gmail.com*

DISEÑO Y COMPOSICIÓN: Carmen Chincoa: *produccionRSEL@gmail.com*

ÍNDICE 55/2 (2025)

ARTÍCULOS

Presuposición de alternativas y partículas de foco: el caso de hasta 9

Presupposition of Alternatives and Focal Particles: The Case of hasta 'even'

CARMEN CONTI

Estudio piloto de los familectos en español: optimización metodológica

e implicaciones para el análisis de la variación sociolingüística 37

Pilot Study on Spanish Familects:

Methodological Optimization and Implications for Sociolinguistics

MARÍA ALCEDO-ORTIZ, BÁRBARA MARQUETA-GRACIA Y ANTONIO BENÍTEZ-BURRACO

El estilo nominal en el contexto de la lecturabilidad: una revisión empírica 71

The Nominal Style in the Context of Readability: An Empirical Review

SUSANA AZPIAZU TORRES

La incorporación nominal y la composición nominal en griego clásico:

νομοθέτης como caso de estudio 115

Noun Incorporation and Nominal Compounding in Classical Greek:

Νομοθέτης as a Case of Study

DANIEL AYORA

Generic Null Objects in Spanish 131

Los objetos tácitos genéricos en español

CARLOS MARTÍNEZ-GARCÍA

Diversidad y sofisticación de colocaciones

en la producción escrita de aprendices y hablantes nativos de español 169

Diversity and Sophistication of Collocations in Written Production

of Learners and Native Speakers of Spanish

ROCÍO CUBEROS-VICENTE Y ELISA ROSADO

RESEÑAS 195

ARTÍCULOS



PRESUPOSICIÓN DE ALTERNATIVAS Y PARTÍCULAS DE FOCO: EL CASO DE *HASTA*

Carmen CONTI JIMÉNEZ¹

Universidad de Jaén

Resumen

En este artículo se estudian algunas de las condiciones gramaticales y contextuales que permiten satisfacer la presuposición de alternativas asociadas al segmento focalizado con *hasta*. En la bibliografía especializada *hasta* focal se ha relacionado con la evocación o la interpretación de alternativas, pero no se ha estudiado el comportamiento de *hasta* en tanto que inductor presuposicional ni las condiciones que permiten satisfacer la presuposición de alternativas cuando se emplea esta partícula con distintos tipos de alcance. Este artículo se propone revisar esta cuestión a partir del análisis de *hasta* focal en dos muestras, una oral y otra escrita, formadas por ejemplos del español europeo extraídos de CORPES XXI (versión 0.99). En estas muestras se documentan ejemplos de *hasta* en los que la presuposición de alternativas se bloquea o no se satisface, pero se mantiene la interpretación del enunciado. En estos ejemplos se observan pautas comunes que parecen ser especialmente importantes, como son el tipo de alcance de la partícula, principalmente oracional, y la ausencia de alternativas mencionadas en el discurso inmediato, asociada sobre todo a textos orales no planificados. Aunque en este artículo no se pretende explicar estos casos, sino tan solo dar cuenta de ellos y de las condiciones que parecen favorecerlos, se postulan dos posibles explicaciones de las correlaciones observadas: de una parte, la naturaleza escalar y evaluativa de la partícula, que se superpone a la naturaleza presuposicional; y de otro, el hecho de que el alcance sea oracional, ya que la oración es un tipo de segmento que plantea más dificultades de comutación.

Palabras clave: presuposición de alternativas; adverbios de foco; *hasta*.

1. cconti@ujaen.es;  <https://orcid.org/0000-0003-3075-3575>.

PRESUPPOSITION OF ALTERNATIVES AND FOCAL PARTICLES: THE CASE OF *HASTA* 'EVEN'

Abstract

This paper deals with some of the problems that focal *hasta* 'even' arises for the presupposition of alternatives associated with the focus. Although there are relevant studies of focal *hasta* that relate the occurrence of this particle with alternatives of the focused segment, the scholars have not paid attention to those factors that allow the speakers to satisfy (or not) the presupposition of alternatives. In order to study to what extent *hasta* allows interpreting the existence of alternatives of the focus, this paper is based on the analysis of examples from CORPES XXI (0.99 version). As will be shown, these samples attest cases of *hasta* that fails in the association of the focus with its possible alternatives. We will posit that these examples share some crucial properties, such as the sentential scope of the particle and the lack of alternatives mentioned in the immediate discourse, which is more frequent in oral texts. The explanation of these patterns is outside the scope of this paper, although we will point to the possibility that the semantics, the pragmatics, and the grammar of the particle play a crucial role. On the one hand, *hasta* has a complex semantics allowing that, in the case that the presupposition of alternatives is not satisfied, the presupposed information accommodates to the scalar-evaluative implicature, which is also induced by *hasta*. On the other hand, sentences under the scope of the particle are also complex types that do not easily allow establishing semantic associations with equivalent alternatives, even more if some of these alternatives are not mentioned in the immediate discourse.

Keywords: presupposition of alternatives; focus particles; *hasta*.

RECIBIDO: 07/01/2025

APROBADO: 20/05/2025

1. INTRODUCCIÓN

Las partículas focales, como *hasta* en español (p. ej. *Habló hasta su hermana*), se caracterizan por destacar el segmento sobre el que tienen alcance. A partir de los trabajos de semántica lógica de Rooth (1985, 1992), no es infrecuente asociar la presencia de un foco en la oración con la evocación o la interpretación de alternativas, si bien la forma en la que se define este fenómeno, así como la naturaleza del mismo (semántica y/o pragmática), varía de unos autores a otros (Fox y Katzir, 2011).

En concreto, para Rooth (1985, 1992), la presuposición de alternativas es un efecto semántico asociado a la presencia de un foco. Este efecto se ha estudiado con detalle en los procesos de focalización de lenguas como el inglés o el alemán,

con especial atención al papel desempeñado por los activadores (*triggers*) de presuposición (como las partículas focales inclusivas y excluyentes), y tan solo ha empezado a estudiarse desde hace unos años para algunas partículas del español dentro de la lingüística experimental (v. § 3).

Este artículo se centra en el estudio de *hasta* focal en español europeo, ya que, como trataremos con detalle en la sección 2, esta partícula se ha relacionado en la bibliografía previa con la interpretación, evocación o implicación de alternativas, bien debido a su naturaleza focal (presuposición de alternativas asociadas al foco), bien debido a su naturaleza escalar (implicatura de escalaridad, u ordenación de la entidad en una escala en relación con otros elementos).

Como se mostrará en la sección 3, *hasta* se puede considerar, al igual que las partículas focales aditivas de otras lenguas, un activador de presuposición de existencia de alternativas, como se ilustra en el ejemplo de (1):

- (1) Hasta [en eso]_F era machista la sociedad. (CORPES XXI, escrito)

La oración de (1) expresa dos tipos de contenidos: el aseverado (la sociedad era machista en eso) y el presupuesto (la sociedad era machista en otros aspectos, que no se mencionan en el enunciado). La información presupuesta debe poder satisfacerse para que la oración que contiene el activador presuposicional, como es el caso de *hasta* en el ejemplo, sea interpretable. El modo en el que se satisface una presuposición de alternativas, como se mostrará en la sección 3, guarda una relación estrecha con tres factores: la posibilidad de commutar el segmento focalizado con esas alternativas, la selección de posibles alternativas en el contexto y la mención de algunas de esas alternativas en el discurso inmediato. En el ejemplo de (1), la oración está fuera de contexto discursivo, pero nuestro conocimiento del mundo nos permite idear posibles alternativas para el foco, siempre y cuando sean commutables (pertenezcan al mismo tipo semántico-sintáctico) y guarden algún tipo de relación semántica. De existir un contexto previo, será relevante tener en cuenta, además, qué alternativas se han mencionado y qué alternativas, de no estar mencionadas, son posibles, como veremos en la sección 3.

Sin embargo, la presuposición inducida por *hasta* no siempre se puede satisfacer semánticamente. Así sucede, por ejemplo, en la oración de (2), en la que no cabe interpretar alternativas por las que commutar el segmento focalizado que sean adecuadas desde el punto de vista formal, semántico y contextual:

- (2) Bueno... en eso de distinto hasta [se pasaba tal vez]_{F*} (CORPES XXI, escrito)

En este artículo demostraremos que la satisfacción de la presuposición de alternativas con *hasta* focal depende en gran medida del tipo de segmento sobre el que tiene alcance la partícula y de la mención de alternativas en el discurso inmediato, como veremos en la sección 5.

Para el estudio del comportamiento de *hasta* focal como activador presuposicional, en este artículo se ha realizado una labor de clasificación y análisis de ejemplos que ilustran el uso real de la partícula. En concreto, se han elaborado dos muestras, una oral y otra escrita, de ejemplos extraídos del subcorpus de España de CORPES XXI (versión 0.99), de cuyos criterios de formación se hablará con detalle en la sección 4.1. Estos ejemplos se han clasificado a partir de la posición de la partícula y del tipo de segmento sobre el que tiene alcance (v. § 4.2). A continuación, se han vuelto a clasificar los ejemplos en dos subgrupos: aquellos que permiten satisfacer la presuposición de existencia de alternativas y aquellos en los que esto no es posible. Para el análisis de cada subgrupo, se han tenido en cuenta dos criterios: el tipo sintáctico de segmento sobre el que tiene alcance la partícula y la mención o no de alternativas en el discurso inmediato (v. § 5). El análisis de los datos permite observar que hay dos situaciones en las que, con más frecuencia que en el resto, no se puede satisfacer la presuposición de alternativas para *hasta*: cuando la partícula tiene alcance oracional y cuando no se hace mención en el discurso inmediato a oraciones que se puedan considerar alternativas del segmento focalizado.

El análisis de los datos plantea una serie de interrogantes en torno a dos cuestiones: por qué el alcance oracional favorece que no se satisfaga la información presuposicional y por qué el enunciado en el que esto sucede sigue siendo interpretable. En la sección 6, explicaremos de forma provisional estas dos particularidades, que merecen en el futuro una mayor atención.

Para alcanzar nuestros objetivos, este artículo se ha estructurado en seis secciones, además de la introducción. En la sección 2, se presentan las propiedades sintácticas y pragmático-discursivas de *hasta* focal señaladas en la bibliografía previa. En la sección 3, se revisa el concepto de presuposición de alternativas ligado a la focalización y se presentan algunas de las propuestas en torno a los factores que permiten satisfacer la presuposición de alternativas. En la sección 4, se clasifican e ilustran los casos de *hasta* focal para el español europeo actual extraídos de CORPES XXI (versión 0.99), atendiendo a la posición y al tipo de alcance de la partícula. En la sección 5, se revisan los ejemplos de las dos muestras según permitan o no satisfacer la presuposición de alternativas para el foco, prestando especial atención al tipo de alcance y a la mención de alternativas en el discurso

inmediato. En la sección 6 se analizan los datos obtenidos y, en la sección 7, se ofrecen las conclusiones de este artículo.

2. LA PARTÍCULA FOCAL *HASTA* EN LOS ESTUDIOS PREVIOS

Hasta se cataloga en la actualidad como preposición o como adverbio de foco. Se considera preposición cuando expresa límite temporal (3a)², espacial (3b) o de cantidad (3c), y adverbio de foco en los casos en los que la partícula destaca el segmento sobre el que tiene alcance (4)³.

- (3) a. Abrimos hasta las tres.
- b. Tienes que caminar hasta el muro.
- c. Puedo gastarme hasta sesenta euros.
- (4) En ese momento resultaba incómodo hasta para él. (CORPES XXI, escrito)

Dentro de los adverbios de foco, *hasta* forma parte de los denominados aditivos o de inclusión (Sánchez, 1999; RAE y ASALE, 2025), junto a *aun*, *aún*, *inclusive*, *incluso*, *también* y *todavía* (RAE y ASALE, 2025, p. 3643). Estos adverbios añaden el foco dentro de un conjunto de elementos, que pueden estar expresados en el enunciado o en el contexto discursivo previo, o bien pueden quedar implícitos⁴.

Como adverbio de inclusión, *hasta* focal responde a las propiedades que König (1991) atribuye a las partículas aditivas. En concreto, la presencia de *hasta* no anula

2. Como preposición temporal, *hasta* puede expresar límite máximo (*Abrimos hasta las tres*) y también punto de partida, en cuyo caso se exige negación en la mayoría de las variedades del español (*No abrimos hasta las tres*). En México, América Central, Colombia y zona costera de Ecuador, la preposición *hasta* tiene ambos sentidos sin necesidad de negación. Por ejemplo, en *Abrimos hasta las tres*, *hasta* puede significar las dos cosas (*Diccionario panhispánico de dudas*). Véanse Dominicic (1982) y Miyoshi (2015) para estos usos.

3. En México, *hasta* adverbio tiene también valor enfático en *hasta atrás* y *hasta adelante* ('lo más atrás/adelante posible') (*Diccionario panhispánico de dudas*). Véase Miyoshi (2015) para los valores de este adverbio en el español de México. Cabe señalar que la clasificación de *hasta* no es uniforme en CORPES XXI (versión 0.99). *Hasta* focal se considera preposición, en lugar de adverbio, en muchos de los casos en los que el foco es un SN (p. ej. *Hasta miedo me dan*).

4. Las partículas *tampoco*, *ni siquiera* y *ni aun* se clasifican dentro de los adverbios de foco de inclusión en RAE y ASALE (2025, p. 3643), ya que se consideran variantes negativas de *también*, *incluso* y *aun*, respectivamente. Salazar & Carretero (2025) consideran *tampoco* un caso de negación aditiva. Para Loureda Cruz, Recio y Rudka. (2021, p. 142), *ni siquiera* y *ni aun* son, sin embargo, adverbios de exclusión, junto a *solo* y *ni*. Los adverbios de exclusión en RAE y ASALE (2025, p. 3643) son *solo*, *tan solo*, *solamente*, *únicamente*, *exclusivamente*, *no más* y *nada más*. Dado que este asunto no forma parte de nuestros objetivos, dejaremos abiertas las distintas opciones de clasificación, ya que no afectan al estudio de *hasta*. En cualquier caso, al listado de uno u otro tipo debería añadirse el adverbio *jamás*, que es partícula focal.

las condiciones de verdad de la correspondiente oración sin la partícula; dota de fuerza cuantitativa al significado de la oración; y expresa el extremo de una escala formada por alternativas (König, 1991, p. 33). Por ejemplo, en una oración como *Habló hasta su hermana*, es cierto que habló su hermana; presuponemos que, además de su hermana, hablaron otras personas; e interpretamos que el emisor del mensaje considera poco probable que su hermana hablara (es decir, del conjunto de personas que hablaron, su hermana era, en la escala de probabilidades, la menos esperada).

Hasta expresa también un contenido escalar, pues sitúa el elemento del foco dentro de una escala. Además de *hasta*, dentro de los adverbios de inclusión se consideran escalares *aun*, *incluso* y *ni siquiera* (RAE y ASALE, 2025, p. 3660), a los que Loureda, Cruz, Recio y Rudka (2021) añaden el adverbio *también*, pues estos autores parten de la noción de escala de Portolés (2007, 2009). En concreto, Portolés (2007) considera que las escalas aditivas pueden ser no culminativas, si el último elemento no se considera más informativo que el anterior o los anteriores (como sucede con *además* y *también*); o pueden ser culminativas, si el elemento añadido es más informativo que los anteriores (como sucede con *hasta*). Por ejemplo, en el caso de *también*, se añade un elemento que se considera el último de una enumeración, pero que no ocupa un extremo de la escala. Así, en *Luisa compró melón y también plátano*, el elemento añadido no forma parte de una lista de elementos jerarquizados. Sin embargo, en el caso de *hasta*, el elemento añadido ocupa un extremo de la escala y su valor informativo es, por tanto, distinto respecto del resto de elementos que la componen, y que se consideran sus alternativas.

En este artículo separaremos por conveniencia explicativa el valor aditivo del escalar (*hasta* añade un elemento a un conjunto o lo ordena en una escala) y tendremos en cuenta, además, el valor evaluativo que se ha atribuido a la partícula *hasta* en la bibliografía previa. En particular, Sánchez (1999, p. 1105) considera que *hasta* es un cuantificador presuposicional, pues induce «la interpretación cuantitativa del elemento al que modifican por implicación de existencia (o inexistencia) de otros elementos». La autora añade que estos cuantificadores presuposicionales pueden tener un «matiz valorativo o evaluativo», al igual que los cuantificadores de grado (Sánchez, 1999, p. 1107).

Albelda (2014), por su parte, considera que *hasta* desempeña una única función pragmática general, que denomina *de intensificación* y que divide en dos propiedades: la escalaridad y la evaluación. En opinión de esta autora, *hasta* es escalar porque permite ordenar el elemento sobre el que tiene alcance en relación con otros elementos que pueden estar presentes o no estarlo; y es evaluativa porque

el empleo de esta partícula permite procesar una instrucción por la que se indica que la información que sigue es la que el hablante juzga como más extrema y menos esperada.

En otro orden de cosas, cabe destacar que los adverbios de foco han suscitado cierto debate en torno a su inclusión dentro del grupo más general de las partículas discursivas y, dentro de estas, a su relación con los marcadores discursivos. De acuerdo con Loureda *et al.* (2021, pp. 16-17), las partículas discursivas se caracterizan en conjunto por compartir una serie de propiedades formales, semánticas y pragmáticas, y, en algunos casos, también sintácticas. Desde el punto de vista formal, las partículas discursivas conforman una macroclase que es transcategorial. El conjunto está constituido por elementos categorialmente diversos (p. ej. adverbios, verbos y constituyentes) que son invariables (p. ej. adverbios) o presentan un alto grado de fijación (p. ej. grupos preposicionales o verbos), tal y como señalan, en otros términos, Loureda *et al.* (2021, p. 16). La partícula *hasta* posee las propiedades formales esperables de una partícula discursiva.

Desde el punto de vista semántico, las partículas discursivas se caracterizan por guiar las inferencias que se realizan en la comunicación. Esta propiedad es común a todos los elementos de la macroclase. Así, por ejemplo, *hasta* posee un significado procedimental adscrito a la propia palabra, que nos permite interpretar correctamente el enunciado en el que aparece inserta⁵. De acuerdo con Briz, Pons y Portolés (2008), el hecho de que la interpretación de *hasta* focal dependa, entre otros factores, del discurso previo, del contexto y de nuestro conocimiento del mundo justifica su inclusión dentro de las partículas discursivas⁶.

En lo que respecta a las propiedades sintácticas de estas partículas, como indican Loureda *et al.* (2021, p. 17), existen diferencias importantes que permiten clasificar los elementos de esta macroclase en dos grandes grupos: de una parte, los marcadores del discurso, que son elementos extraoracionales (p. ej. *sin embargo*), y, de otra, las partículas con incidencia dentro de la oración, como es el caso de las partículas focales, que Loureda *et al.* (2021: 18) denominan *operadores*.

En lo que se refiere a las propiedades sintácticas de las partículas focales cabe hacer algunos comentarios de carácter general. Al menos la partícula focal *hasta* no

5. De acuerdo con Loureda *et al.* (2021, p. 282), los elementos con significado procedimental determinan la manera en la que la cognición humana debe tratar la información que proporcionan las unidades con significado conceptual.

6. Para estos autores, *hasta* se define en términos similares a los expuestos hasta ahora en esta sección: «destaca un elemento del discurso como el límite máximo inesperado de una serie de la que también forman parte otros elementos (expresos previamente o, lo que es más frecuente, sobrentendidos)».

forma parte de la representación léxico-semántica del predicado (no aparece en la estructura predicativa) y no forma parte tampoco de la representación semántico-lógica de la oración (su aparición, como vimos más arriba, no cambia los valores de verdad de la oración). *Hasta* focal se encuentra dentro de los límites formales de la oración, presenta restricciones distribucionales importantes y su posición determina la estructura informativa de la oración. Como indican Briz, Pons y Portolés (2008), *hasta* no puede funcionar de forma independiente, se puede combinar con *incluso* (*incluso hasta/hasta incluso*) y precede siempre al miembro discursivo que destaca. De acuerdo con estos autores, *hasta* puede tener alcance sobre toda la oración (v. 5a) o sobre un segmento (un sintagma nominal, v. 5b; adjetival, v. 5c; preposicional, v. 5d; adverbial, v. 5e; o una oración subordinada, v. 5f)⁷:

- (5) a. Eso sería como descubrir que Emilio Aragón fuma porros. Hasta sería interesante.
- b. Cambian o modifican el conocimiento, las creencias, los deseos, la situación mental del que escucha, sus movimientos y hasta su comportamiento.
- c. La respuesta del Zaragoza fue enorme, bella, espectacular y hasta efectiva.
- d. [...] y meter diminutivos hasta en la sopa [...]
- e. [...] ambos son dos ejemplos preclaros de *espíritus* tenazmente —¡hasta imperiosamente!— comunicativos [...]
- f. Miren, de él (de Curro Romero) dicen de él dicen que es torero hasta cuando camina.

En resumen, *hasta* es una partícula focal que pertenece a la clase de los adverbios. Su contenido procedimental se relaciona con tres tipos de instrucciones: aditiva (añádase la entidad del foco a un conjunto de entidades, entendidas como sus alternativas), escalar (entiéndase que esa entidad añadida forma parte de una escala y que esa entidad es la más extrema, por encima de las entidades o alternativas, que están ordenadas por debajo) y evaluativa (entiéndase que el locutor considera sorpresivo o poco probable que esa entidad esté en la escala). En lo que respecta a sus propiedades gramaticales, *hasta* es una partícula intraoracional con restricciones distribucionales: precede inmediatamente al elemento focalizado y, por consiguiente, carece de alcance a distancia.

Como veremos en la sección 4, las propiedades sintácticas señaladas por Briz, Pons y Portolés (2008) se constatan en las muestras obtenidas, además de otras que no se mencionan en el diccionario de estos autores (v. § 4.2). Veremos, sin embargo,

7. Los ejemplos de (5) proceden de Briz, Pons y Portolés (2008).

que la presuposición de alternativas no siempre se satisface cuando empleamos *hasta* y que la posibilidad de satisfacerla parece depender de factores tales como el tipo de segmento sobre el que tiene alcance la partícula y la mención o no de alternativas en el discurso inmediato (v. § 5).

3. LAS PARTÍCULAS FOCALES Y LA PRESUPOSICIÓN DE ALTERNATIVAS

La presuposición de alternativas se ha relacionado en la bibliografía previa con distintos procedimientos de focalización⁸, entre los que destacan el empleo de mecanismos prosódicos, como se ilustra en (6a), y el empleo de partículas, como en (6b):

- (6) a. Compró [MELÓN]_F.
 b. Compró también [melón]_F.

En las oraciones de (6), es cierto que se compró melón y se presuponen alternativas al foco que no están mencionadas en el enunciado. En (6b) presuponemos que existen otras cosas que se compraron y, en (6a), presuponemos que existen cosas que no se compraron.

La aparición de la partícula *hasta* en una oración también da lugar a la presuposición de existencia de alternativas, como se indicó en la introducción de este artículo. En *Compró hasta [melón]_F*, es cierto que se compró melón y presuponemos que se compraron otras cosas. En esta oración hay, por tanto, dos contenidos: el aseverado (compró melón) y el presupuesto (compró otras cosas, como peras, manzanas, naranjas, etc.). *Hasta* es, por tanto, un inductor o activador presuposicional, como demuestra además el hecho de que la presuposición que induce no se vea afectada por la presencia en el enunciado de marcadores de estatus (modalidad epistémica) o de fuerza ilocutiva (interrogación), o por la integración del enunciado en una estructura condicional⁹. Así, por ejemplo, en *Quizá compró hasta melón, ¿Compró hasta melón? y Si compró hasta melón, estará esperando visita*, la presuposición de que se compraron otras cosas queda fuera del alcance de esos marcadores¹⁰.

8. Véanse, entre otros, Rooth (1985, 1992) y Gotzner (2017).

9. Para estas pruebas, véase el trabajo de van der Sandt (2012, p. 330). *Hasta* no es indiferente, sin embargo, a la prueba de la negación, ya que es una partícula aditiva afirmativa (p. ej. **No compró hasta melón*).

10. La presuposición de alternativas también se ha relacionado con el empleo de expresiones escalares bajo la denominada implicatura de escalaridad (Fox y Katzir, 2011). Se ha de tener en cuenta que *hasta* es una partícula escalar, además de aditiva.

La presuposición de alternativas se ha estudiado con detalle para distintos tipos de focalización en inglés y en alemán dentro de la semántica lógica y la semántica cognitiva experimental. El punto de partida de estos estudios se encuentra en los trabajos de Rooth (1985, 1992), quien enuncia la denominada *asociación del foco*. Según este autor, la focalización tiene como efecto semántico la presuposición de alternativas por las que se puede comutar el segmento focalizado. Esta última propiedad es la que los especialistas denominan *comutación*¹¹, consistente en la posibilidad de reemplazar el segmento focalizado por una alternativa que pertenece al mismo tipo semántico (Rooth, 1985, 1992) y muestra una complejidad equiparable (Fox y Katzir, 2011). En una oración como *Adela limpió también los cristales del coche*, presuponemos alternativas que pertenecen al mismo tipo semántico que el segmento focalizado (p. ej. un tipo de entidad) y que poseen una complejidad estructural similar (p. ej. un sintagma nominal). Una alternativa posible a *Adela limpió también los cristales del coche* podría ser *Adela limpió las ruedas del coche*, ya que *las ruedas* mantiene la propiedad de comutación con el segmento focalizado: pertenece al mismo tipo semántico (denotación de individuos) y posee la misma complejidad estructural (es un SN). La propiedad de comutación explicaría por qué no podemos presuponer que otras acciones de limpiar, como *Adela limpió con lejía*, son alternativas posibles de *Adela limpió también los cristales del coche*, puesto que no se puede establecer una correlación con un elemento comutable para *los cristales*.

De acuerdo con Rooth (1985, 1992), la propiedad de comutación es formal, pero la selección de las alternativas posibles está condicionada por factores contextuales. Actualmente, existen numerosas propuestas a la hora de caracterizar los procesos que tienen lugar en la selección de las alternativas¹². Así, por ejemplo, se admite por lo general que las alternativas han de estar relacionadas semántica o conceptualmente, pero no hay acuerdo a la hora de determinar el tipo de procedimiento de selección que parece tener lugar. Si volvemos al ejemplo anterior, *Adela limpió también los cristales del coche*, observamos que, cuando presuponemos alternativas, estamos realizando una selección entre un número considerable de alternativas posibles. Se acepta habitualmente que los hablantes activamos y desactivamos alternativas según distintos criterios semánticos (p. ej. asociaciones semánticas) y según la

11. Para estas cuestiones, puede verse el trabajo de Fox y Katzir (2011).

12. Véase también Fox y Katzir (2011) a este respecto.

situación descrita en el contexto¹³. En el ejemplo anterior, el foco *los cristales del coche* activa un conjunto de alternativas posibles relacionadas semánticamente en distintos planos (tanto con las partes de un coche como con lo que puede evocar en sí mismo un coche), de ahí que sea necesario seleccionar aquellas asociaciones que encajan tanto en la semántica oracional (entidades asociadas a un coche que sean lavables, por ejemplo) como en el contexto discursivo. De hecho, en el ejemplo citado, parece más probable presuponer alternativas asociadas semánticamente a las partes de un coche que a entidades susceptibles de ser limpiadas, pero que no guarden relación con las partes de un vehículo.

El contexto discursivo también determina la selección de las alternativas. De acuerdo con Kim, Gumlogson, Tanenhaus y Runner (2015), un ejemplo fuera de contexto, como *Jane only has some apples* ('Jane solo tiene algunas manzanas'), permite presuponer unas alternativas de *manzana* que son semánticamente afines, como peras, naranjas, etc. Sin embargo, en el siguiente contexto discursivo ofrecido por los autores, las alternativas para ese mismo enunciado son muy distintas (traducimos del original inglés al español): *Jane está sola en la cocina por la noche cuando oye a un ladrón intentando entrar. Mira a su alrededor buscando algo que lanzarle cuando salte por la ventana. Jane solo tiene algunas manzanas* (Kim, Gumlogson, Tanenhaus y Runner, 2015, p. 31). Las alternativas para el foco están ahora condicionadas por la situación descrita en el discurso previo, de modo que son alternativas de *manzanas* entidades como cestas, cacerolas, etc.

Parece, en suma, que las alternativas del foco han de ser formalmente commutables (han de pertenecer al mismo tipo semántico lógico), han de estar asociadas semánticamente al foco (dentro de una misma categoría o clase semántica) y han de ser contextualmente apropiadas. Así pues, las alternativas han de satisfacer una serie de condiciones, algunas de las cuales se pueden considerar intraoracionales (commutación) y otras, extraoracionales (asociaciones semánticas y restricciones contextuales).

En los trabajos sobre alternativas del foco también se ha estudiado el grado de relevancia o *saliencia* de las alternativas en oposición al segmento focalizado, sin que tampoco exista consenso en este punto. Para Calhoun (2009), quien realiza un estudio de corpus del foco prosódico en inglés, las alternativas con las que contrasta el foco serán más salientes cuanto más prominente sea en el discurso el

13. Por ejemplo, para Husband y Ferreira (2016), el proceso de selección que se aplica al conjunto de alternativas posibles del foco es similar al que tiene lugar cuando en un texto aparece una palabra polisémica. Ese proceso de selección activa las alternativas semánticamente próximas y desactiva las que no lo son dentro de la proposición focalizada (Husband y Ferreira, 2016, p. 2020).

elemento focalizado. En cambio, Gotzner, Wartenburger y Spalek (2016) observan diferencias entre el empleo del foco prosódico y el empleo de las partículas del alemán *nur* 'solo' y *sogar* 'incluso'. En concreto, las autoras defienden que las partículas conllevan, frente al foco prosódico, una mayor prominencia de las alternativas y una mayor competencia entre estas y el segmento focalizado. Como resultado de esta competencia, el empleo de partículas interfiere en el reconocimiento de las alternativas que aparecen mencionadas, en lugar de facilitarlo (Gotzner, Wartenburger y Spalek, 2016, p. 77). El estudio de estas autoras está basado en tres experimentos con estímulos auditivos, en los que miden el tiempo de reacción de los participantes, que tienen que decidir si una palabra ha aparecido o no en la historia que han escuchado¹⁴.

En el caso del español, también se ha estudiado el efecto de las partículas focales sobre el procesamiento del enunciado (Torres 2020, Loureda *et al.* 2021), si bien las conclusiones a las que se ha llegado a este respecto no coinciden con los estudios del alemán. En concreto, Torres (2020), quien estudia el esfuerzo de procesamiento a través del movimiento ocular en la lectura de enunciados con *hasta*, concluye que la presencia de esta partícula requiere un mayor procesamiento semántico del enunciado que el que requiere su correspondiente enunciado sin partícula, pero no supone, sin embargo, un mayor esfuerzo cognitivo de procesamiento. A partir de los experimentos de Torres (2020), Loureda *et al.* (2021, pp. 146-152) defienden que la aparición de *hasta* en un enunciado satisface los que denominan Principios I y II de la marcación discursiva: el procesamiento del enunciado con *hasta* es distinto del procesamiento del enunciado sin marca (Principio I de la marcación discursiva), pero el esfuerzo de procesamiento del enunciado con *hasta* no puede ser mayor que el del enunciado no marcado (Principio II), es decir, la presencia

14. Las autoras distinguen para ello entre alternativas mencionadas, alternativas relacionadas, pero no mencionadas, y falsas alternativas. Los textos empleados en uno de los experimentos son los siguientes, traducidos al español (Gotzner, Wartenburger y Spalek, 2016, p. 70):

Oración contexto (hablante 1):

<i>In der Obstschüssel</i>	<i>liegen</i>	<i>Pfirsiche, Kirschen und Bananen.</i>
'En el frutero	hay	melocotones, cerezas y plátanos'.

Oración de continuación (hablante 1):

<i>Ich wette, Carsten hat</i>	<i>Kirschen und Bananen gegessen.</i>
'Creo que Carsten comió cerezas y plátanos'.	

Oraciones críticas (hablante 2):

(a) *Nein, er hat nur [Pfirsiche]F gegessen.*

(b) *Nein, er sogar [Pfirsiche]F gegessen.*

(c) *Nein, er hat [Pfirsiche]F gegessen.*

'No, ha comido (a) solo/(b) incluso/(c) [melocotones]F'.

de *hasta* no puede ralentizar el procesamiento. Según esto, el procesamiento de un enunciado como *María compró cerveza, refrescos y hasta cava* no puede verse dificultado por la presencia de la partícula *y*, en ningún caso, puede requerir más esfuerzo cognitivo que el procesamiento del correspondiente enunciado sin partícula (p. ej. *María compró cerveza, refrescos y cava*). En opinión de Loureda *et al.* (2021), *hasta* satisface estos principios independientemente de que haya alternativas codificadas (p. ej. *Compró cerveza y hasta cava; Compró cerveza, refrescos y hasta cava*) o no las haya (p. ej. *Compró hasta cava*).

El trabajo de Torres (2020), el más completo sobre el procesamiento de *hasta*, se basa en enunciados en los que la partícula tiene alcance sobre sintagmas nominales (p. ej. *Alba y Lucas tienen coches y hasta yates*), al igual que sucede en los estudios experimentales del inglés o del alemán que hemos citado con anterioridad. Sin embargo, *hasta* puede tener alcance sobre distintos tipos de segmentos sintácticos, además de grupos nominales, y no todos permiten satisfacer la presuposición de alternativas. Tal y como señala Portolés (2011) a propósito de las partículas aditivas, que, en opinión del autor, se caracterizan por destacar un estado mental menos esperable (foco) que otro (alternativa), «en ocasiones es difícil reconstruir en qué consistiría el estado mental de la alternativa» (Portolés, 2011, p. 8). No por casualidad el ejemplo que lleva al autor a hacer esta reflexión contiene la partícula *hasta* con alcance sobre un sintagma adjetival (*No es que mi santo mi marido sea Carlos Latre, pero cuando las mujeres le ríen las gracias se pone, para mi gusto, hasta patoso*), del que Portolés (2011, p. 8) afirma lo siguiente:

[...] el adverbio de foco *hasta* destaca *patoso* y lo muestra como menos esperable que una alternativa para una fuente que se ha de saturar pragmáticamente; ahora bien, la alternativa tiene un contenido difícil de determinar.

Como indicamos en la introducción, en este artículo nos proponemos estudiar las condiciones que permiten satisfacer la presuposición de alternativas que activa *hasta* focal *y*, por contraste, aquellas otras en las que esto no es posible. Para ello, hemos analizado *hasta* en ejemplos procedentes de textos en los que se hace un uso real de la lengua. En la siguiente sección presentaremos los criterios de selección de los ejemplos (v. § 4.1) y los describiremos teniendo en cuenta la posición de *hasta* y el tipo sintáctico de segmento sobre el que tiene alcance (v. § 4.2). En la sección 5, clasificaremos los ejemplos según permitan o no satisfacer la presuposición de alternativas para el foco. En uno y otro caso tendremos en cuenta el tipo de segmento sobre el que tiene alcance la partícula, la mención de alternativas en el

discurso inmediato (normalmente, previo) y el medio (oral o escrito) de los textos de los que se han extraído los ejemplos.

4. *HASTA EN LA MUESTRA DE EJEMPLOS: POSICIÓN DE LA PARTÍCULA Y SEGMENTO SOBRE EL QUE TIENE ALCANCE*

4.1. *Presentación de la muestra de ejemplos*

En este artículo, *hasta* focal se ha estudiado en dos muestras de ejemplos extraídos de CORPES XXI (versión 0.99): una muestra de 72 ejemplos de textos orales, que contiene todos los ejemplos de *hasta* focal del subcorpus España, y una muestra de 84 ejemplos procedentes de textos escritos. Los ejemplos de la muestra escrita proceden, como los de la muestra oral, del subcorpus España.

Los ejemplos de ambas muestras se obtuvieron al preguntar por *hasta* adverbio en el buscador de CORPES XXI (versión 0.99), ya que este es el estatus categorial de la partícula focal en la bibliografía reciente¹⁵.

Como se ha indicado más arriba, la muestra oral está constituida por todos los ejemplos en los que se documenta *hasta* con valor focal. En el caso de la muestra escrita, los ejemplos se seleccionaron por orden de aparición en los resultados de la búsqueda, tras descartar aquellos usos de *hasta* que no se pueden considerar focales¹⁶.

Las muestras empleadas para el análisis de *hasta* en este artículo son relativamente pequeñas, pero permiten realizar una aproximación a la complejidad de uso de esta partícula. Al ofrecer un número representativo de ejemplos de uso real de *hasta*, las muestras permiten documentar la posición de la partícula, sus distintos tipos de alcance (más allá del meramente nominal) y el estudio de la posible repercusión que puede tener el tipo de alcance en la satisfacción de la presuposición de alternativas.

Como se ha indicado anteriormente, la descripción y el análisis de los ejemplos se han realizado a partir de dos criterios gramaticales (v. § 4.2 para más detalles), que no se han tenido en cuenta a la hora de estudiar con anterioridad la presuposición de alternatividad y que creemos son, sin embargo, necesarios: la posición de *hasta* en la oración y el tipo sintáctico de segmento sobre el que tiene alcance la partícula. De los 72 ejemplos de la muestra oral, 35 de ellos ejemplifican casos de *hasta* con alcance sobre sintagmas preposicionales; 32 sobre oraciones; 2 sobre sintagmas adjetivales; y

15. Cabe advertir aquí que los ejemplos de *hasta* focal con alcance sobre sintagmas nominales se obtienen en CORPES XX (versión 0.99) principalmente cuando se pregunta por *hasta* preposicional, no adverbial.

16. Los ejemplos completos de las dos muestras empleadas en este artículo, en una versión preliminar, se pueden consultar en la base de datos FOCAL-ES, de acceso abierto en <<https://focal-es.vercel.app/>>.

1 sobre el predicado. De los 84 ejemplos de la muestra escrita, 51 ejemplifican casos de *hasta* con alcance sobre oraciones; 25 sobre sintagmas preposicionales; 3 sobre el predicado; 1 sobre sintagma adverbial; 1 sobre sintagma adjetival; 1 sobre sintagma nominal; 1 sobre subordinada; y 1 sobre cláusula de gerundio.

En lo que se refiere a la posición de *hasta* y al tipo de alcance de la partícula, no se observan diferencias significativas entre las dos muestras citadas. Por este motivo, en la sección 4.2 emplearemos indistintamente ejemplos procedentes de una y otra. Sin embargo, como veremos en la sección 5, el hecho de que el discurso esté o no planificado sí parece condicionar el número de ejemplos en los que no se puede presuponer alternatividad para el segmento focalizado, que es mayor en la muestra de ejemplos procedentes de textos orales.

4.2. *La posición de hasta focal y los tipos de segmentos sobre los que tiene alcance*

El estudio de la muestra de ejemplos constata las pautas distribucionales y los tipos de alcance de *hasta* focal que ya señalaron Briz, Pons y Portolés (2008). *Hasta* precede siempre al constituyente que destaca (o sobre el que tiene alcance), se puede combinar con *incluso* y puede tener alcance sobre toda la oración o sobre una parte de ella¹⁷. Se observan, en cambio, otras pautas no señaladas con anterioridad. Además de los tipos sintagmáticos documentados por Briz, Pons y Portolés (2008), *hasta* focal aparece en la muestra con alcance sobre el predicado de la oración (SV), sobre cláusulas de verbo no finito y sobre partes de perífrasis o predicados complejos que no son analizables como estructuras sintagmáticas.

4.2.1. *Con alcance sobre oraciones*

Consideramos que *hasta* tiene alcance sobre oraciones cuando ocupa la primera posición del enunciado y hay sujeto omitido (7a) o el sujeto explícito es posverbal (7b):

- (7) a. Va diciendo por ahí que es mi novio, hasta [me obliga a acompañarle los domingos por la tarde]. (CORPES XXI, escrito)
- b. yo creo que hasta [podría ir tu padre]. (CORPES XXI, escrito)

Como se ilustra en (7a), *hasta* puede aparecer en posición inicial absoluta, aunque lo más habitual en la muestra es que *hasta* esté precedido por una conjunción

17. Para una descripción análoga de *inclusive* e *incluso*, partículas estrechamente relacionadas con *hasta*, puede consultarse el trabajo de Felú y Pato (2022).

subordinante (7b y 8a), una conjunción coordinante (8b) o un modificador/adjunto extraoracional (9):

- (8) a. que yo de hecho confieso que hasta [tengo problemas para leer cosas que han escrito otros]. (CORPES XXI, oral)
- b. Se pochan y hasta [huelen mal]. (CORPES XXI, escrito)
- (9) a. el año pasado ha hasta [hubo un momento que me parecía que podía haber ganado la liga]. (CORPES XXI, oral)
- b. a lo mejor hasta [le di clase yo]. (CORPES XXI, escrito)

Cabe señalar por último que *hasta* puede tener alcance sobre oraciones negadas (10). La interpretación del valor o valores de *hasta* en estos casos es opaca o poco transparente:

- (10) a. y hasta [no te hubiera importado tener un accidente]. (CORPES XXI, escrito, España)
- b. ya sabe que no voy a hacer comentarios sobre los responsables políticos // no lo voy a hacer sobre lo que han dicho unos y otros / algunas incluso hasta [no las conozco en concreto]. (CORPES XXI, oral, España)

4.2.2. *Con alcance estrecho*

Cuando *hasta* focal tiene alcance sobre sintagmas preposicionales o nominales, estos pueden aparecer en posición canónica (11) o antepuestos (12):

- (11) a. Tenía hasta [visiones]. (CORPES XXI, escrito)
- b. pero bueno mira a la hora de la verdad acabaremos entrevistando hasta [a tu perro]. (CORPES XXI, oral)
- (12) a. Hasta [miedo] me dan. (CORPES XXI, escrito)
- b. Hasta [en eso] era machista la sociedad. (CORPES XXI, escrito)

En los casos en los que el elemento focalizado se antepone, el sujeto explícito ocupa una posición posverbal (v. *la sociedad* en 12b), tal y como sucede en otras estructuras de anteposición en español europeo (p. ej. *A MARÍA vio Luis*) o se omite (12a).

Como se ha indicado en la introducción, *hasta* puede tener alcance sobre otros tipos de constituyentes, pero estos se documentan en menos ejemplos. Como se

ilustra a continuación, la partícula puede tener alcance sobre sintagmas adjetivales (13) y sobre sintagmas adverbiales (14):

- (13) O como Gonzalo el policía, tan educado, hasta [tan cortado]. (CORPES XXI, escrito)
- (14) Me siento hasta [mal] y con la tripa revuelta. (CORPES XXI, escrito)

Para este tipo de alcance no se documentan ejemplos de anteposición en las muestras, si bien son posibles en español (p. ej. *Hasta mal me sentó*).

Hasta también puede tener alcance sobre subordinadas, como señalaron Briz, Pons y Portolés (2008), tanto con verbo finito (15) como con verbo no finito (cláusulas de infinitivo, v. 16a, y de gerundio, v. 16b):

- (15) Nos han manipulado tanto tantas veces que hasta [de lo que podemos estar orgullosos] lo utilizamos como arma arrojadiza. (CORPES XXI, escrito)
- (16) a. Hasta [comer y dormir] era secundario. (CORPES XXI, escrito)
- b. mientras el agente López tiraba en busca de la familia de ese pobre Romero, casi empujando la ambulancia, tal cual, y hasta [abriéndole paso al enfermo para que [...]] (CORPES XXI, escrito)

Igualmente, en las muestras se documentan ejemplos en los que *hasta* tiene alcance solo sobre el predicado (SV) (v. 17). En tales casos, el sujeto léxico precede a la partícula y queda fuera del foco:

- (17) a. yo hasta [me creí que el niño estaba vivo]. (CORPES XXI, oral)
- b. el pobre Modesto hasta [había dejado viuda]. (CORPES XXI, escrito)

Cabe señalar, por último, que *hasta* puede aparecer en el interior de algunas perifrasis o quasi perifrasis de infinitivo (18) y de gerundio (19). En unos casos, cabe interpretar que *hasta* tiene alcance sobre el verbo auxiliado (v. 18a); en otros, sobre el predicado (verbo auxiliado más complementos) (v. 18b-c y 19):

- (18) a. Con los ojos tapados por el antifaz podría hasta [llorar] sin que nadie se diera cuenta. (CORPES XXI, escrito)
- b. un país que no reconoce el esfuerzo que llega hasta [a entregar la vida] cuando se lucha por la paz y por la justicia. (CORPES XXI, oral)
- c. Pues me ha explicado el infeliz que ha llegado hasta [a pensar en robar] [...] (CORPES XXI, escrito)

- (19) la situación empezaba a superarla y andaba ya hasta [traicionando sus convicciones más profundas sobre una relación]. (CORPES XXI, escrito)

5. HASTA FOCAL Y LA PRESUPOSICIÓN DE ALTERNATIVAS

La clasificación de los ejemplos de las muestras que se ha presentado en la sección anterior se ha tomado como punto de partida para una clasificación posterior de los ejemplos en dos subgrupos: aquellos que permiten presuponer alternativas al foco y aquellos otros que no lo permiten. Los ejemplos que permiten presuponer alternativas para el foco son aquellos en los que el segmento focalizado es fácilmente comutable, al menos, por una alternativa, que ha de ser formal, semántica y contextualmente adecuada; por el contrario, los que no permiten presuponer alternativas son aquellos en los que no se dan estas condiciones¹⁸.

A estos dos subgrupos de ejemplos se han aplicado los siguientes criterios de análisis: (i) la procedencia oral o escrita del texto del que se extrae el ejemplo; (ii) el tipo sintáctico de segmento sobre el que tiene alcance la partícula; y (iii) la mención o no de alternativas en el discurso inmediato¹⁹.

Como veremos en las secciones 5.1 y 5.2, la diferencia entre la procedencia oral o escrita de las muestras empleadas es relevante en el estudio de la presuposición de alternativas, ya que la muestra oral contiene un 23% de ejemplos en los que no se satisface la presuposición de alternativas para el foco, frente a un 4% de ejemplos en la muestra escrita. La explicación de esta diferencia excede los límites de este artículo, pero su existencia justifica que se traten por separado las dos muestras en sendas secciones.

El criterio relativo al tipo de segmento sintáctico sobre el que tiene alcance la partícula no se ha aplicado al estudio de la satisfacción de la presuposición de alternativas asociada al foco en la bibliografía previa. Sin embargo, como mostraremos en la sección 5.2, entre los ejemplos en los que no se satisface la presuposición de alternativas predominan los casos en los que *hasta* tiene alcance sobre oraciones. Esto indica que las oraciones son segmentos más difíciles de comutar que otros, dado que las posibles asociaciones semánticas y contextuales con alternativas son bastante más complejas.

18. La autora de este artículo ha sido la única evaluadora de la presuposición de alternativas de los ejemplos. Sería interesante revisar los resultados obtenidos mediante un experimento, en la línea de lo que se ha hecho para las partículas focales de otras lenguas.

19. Hemos considerado *discurso inmediato* el fragmento discursivo que arroja CORPES XXI (versión 0.99) cuando se obtiene el ejemplo.

El tercer criterio que se ha tenido en cuenta, relativo a la mención en el discurso inmediato de las alternativas, se ha utilizado en numerosos trabajos de carácter experimental en el estudio de la presuposición de alternativas y la focalización, con o sin partículas, como se mostró en la sección 3. En las siguientes secciones, probaremos que, en efecto, la mención de alternativas conmutables por el foco está en correlación directa con la satisfacción de la presuposición de alternativas desencadenada por *hasta* en nuestras muestras de ejemplos.

5.1. Presuposición de alternativas en la muestra oral

En la muestra oral, la mención en el discurso inmediato de alternativas garantiza la satisfacción de la presuposición de existencia de alternativas para el segmento focalizado por *hasta*. Como se ilustra en (20), esas alternativas mencionadas son por lo general miembros de una estructura de coordinación (v. el subrayado en 20a) o de yuxtaposición (v. el subrayado en 20b):

- (20) a. porque la señora Sáez / no habla // en los argumentos / y hasta [en los tiempos] // con las declaraciones que hacen los dirigentes / de un determinado / partido. (CORPES XXI, oral)
- b. aquí lo prioritario es su perfil / su capacidad / sus cualidades humanas / profesionales / técnicas / hasta [de conocimiento de lo que es el servicio exterior]. (CORPES XXI, oral)

Las alternativas mencionadas pertenecen al mismo tipo semántico-lógico que el segmento focalizado, además de desempeñar la misma función sintáctica, como se observa en los ejemplos anteriores de (20).

El tipo de segmento sintáctico sobre el que tiene alcance *hasta* en los casos en los que hay alternativas mencionadas presenta una considerable variedad formal: *hasta* tiene alcance sobre sintagmas preposicionales (60% de ejemplos) (21a), sobre oraciones (30%) (21b), sobre predicados (5%) (21c) y sobre sintagmas nominales (5%) (21d).

- (21) a. La receta de este exquisito plato tradicional varía en cada pueblo // en cada barrio / y hasta [en cada casa]. (CORPES XXI, oral)
- b. si no mira lo que está pasando con Twitter por ejemplo / cómo Twitter / genera debates de opinión y // y a veces hasta [conforma opinión en solo ciento cuarenta caracteres]. (CORPES XXI, oral)

- c. Mucha gente se deprime y sufre / y alguno hasta [se quita la vida] pero no daña. (CORPES XXI, oral)
- d. es el mismo racismo / que se plantea respecto a los napolitanos se plantea hasta incluso hasta hasta [los propios romanos]. (CORPES XXI, oral)

Cuando no hay alternativas mencionadas en el discurso inmediato, la satisfacción de la presuposición de alternatividad se mantiene en algunos ejemplos de la muestra oral, pero no en otros. En los casos en los que se satisface la presuposición de alternativas, la partícula documenta alcance sobre distintos tipos de segmentos sintácticos, si bien predominan los ejemplos con alcance sobre sintagmas preposicionales, como los de (22)²⁰:

- (22) a. yo soy cariñoso hasta [con Bonilla]. (CORPES XXI, oral)
- b. que yo ahora hasta hasta [para hacer quimioterapia] te dan lista de espera. (CORPES XXI, oral)
- c. ahora se habla hasta [del verano]. (CORPES XXI, oral)
- d. porque / yo creo que si son unos profesionales deberían de jugar / hasta [en el campo de aquí del pueblo de arena]. (CORPES XXI, oral)

En los ejemplos de (23) no hay alternativas mencionadas en el discurso inmediato, pero se satisface la presuposición de su existencia: se presupone que el sujeto de (22a) es cariñoso con otras personas, además de con Bonilla; que dan lista de espera para otros tratamientos, además de para la quimioterapia (22b); que se habla de otras cosas, además del verano (23c); y que los profesionales del fútbol deberían jugar en otros lugares, incluido el campo de arena del pueblo (22d). Podemos presuponer alternativas más o menos definidas para el foco de las oraciones de (22) porque el segmento focalizado, la semántica oracional y nuestro conocimiento del mundo nos lo permiten. El segmento focalizado es fácilmente comutable por alternativas semánticamente y contextualmente posibles.

También permiten presuponer alternativas para el foco ejemplos en los que *hasta* tiene alcance sobre formas verbales no finitas en construcciones cuasiperifrásicas, como en (23a), y sobre sintagmas adjetivales, como en (23b):

- (23) a. un país que no reconoce el esfuerzo que llega hasta [a entregar la vida] cuando se lucha por la paz y por la justicia. (CORPES XXI, oral)
- b. resulta hasta [a aburrido] ¿no? (CORPES XXI, oral)

20. En concreto, se documentan un 65% de sintagmas preposicionales, un 26% de oraciones, un 3% de predicados, un 3% de sintagmas nominales y un 3% de sintagmas adjetivales.

Los ejemplos en los que no se satisface la presuposición de alternativas con *hasta* presentan también variedad en el tipo de alcance de la partícula, si bien predominan los ejemplos en los que *hasta* tiene alcance sobre oraciones: en concreto, de 30 ejemplos de *hasta* con alcance oracional, 11 no permiten presuponer alternativas claras para la oración focalizada.

Los ejemplos de (24) y (25) ilustran los distintos tipos de segmentos documentados en los que no se satisface la presuposición de alternativas: (24a) para *hasta* con alcance sobre sintagma preposicional (13% de los casos); (24b) sobre predicados (12%); (24c) sobre sintagma adjetival (6%); y los de (25) para *hasta* con alcance sobre oraciones (69% de los casos).

- (24) a. ya decía yo que estamos hasta [ante un ejemplo especial] y diría que casi único. (CORPES XXI, oral)
 - b. no me refería a Jiménez de Parga / me refería al Partido Socialista Obrero Español / que son cosas distintas // a partir de ahí / di mi opinión mi opinión es que estamos en un sistema de mandato representativo y no de mandato imperativo / que es el sistema que existe en todos los países del mundo // esa es mi opinión y yo hasta [puedo creer que estoy equivocado]. (CORPES XXI, oral)
 - c. Y había una mujer, que eso está hasta [asqueroso] decirlo, que no tiraba la orina por la mañana. (CORPES XXI, oral)
- (25) a. tengo como un gran equipo / el año pasado ha hasta [hubo un momento que me parecía que podía haber ganado la liga]. (CORPES XXI, oral)
 - b. pero después con las otras bien a ver gallego si no fuera por la literatura que hay que estudiar / bien aunque bueno tampoco / como nunca hablo / gallego / después a veces ya hasta [me trabo al hablar gallego]. (CORPES XXI, oral)
 - c. el diablo me parece / un invento que además hasta [creo que lo ha suprimido el Wojtyla]. (CORPES XXI, oral)

De la descripción realizada en esta sección se desprende que, cuando el segmento focalizado es una oración y no hay alternativas mencionadas en el discurso inmediato (p. ej. miembros de estructuras coordinadas o yuxtapuestas), decrecen las posibilidades de satisfacer la presuposición de existencia de alternativas de *hasta* focal.

5.2. Presuposición de alternativas en la muestra escrita

Los ejemplos en los que se satisface la presuposición de alternativas para el foco en la muestra escrita son mayoritarios y presentan pautas similares a los ejemplos

de la muestra oral. Tan solo en los tres ejemplos siguientes es difícil satisfacer la presuposición de alternativas para el foco, en los que, además, *hasta* tiene alcance sobre oraciones:

- (26) a. no con be, de ir para arriba, que también los hay, no te creas, Mayte, y que hasta [suelen ser los mismos], como el doctor siempre dice. (CORPES XXI, escrito)
- b. no sé lo que ganas, pero supongo que hasta [saldrás beneficiado], a lo mejor te devuelven una parte de lo que te han retenido. (CORPES XXI, escrito)
- c. bueno... en eso de distinto hasta [se pasaba tal vez]. (CORPES XXI, escrito)

En el resto de los ejemplos de la muestra escrita se satisface la presuposición de existencia de alternativas. En 44 de los ejemplos analizados, aparecen mencionadas alternativas en el discurso inmediato, en estructuras de coordinación o yuxtaposición. El tipo de alcance de la partícula es variado. *Hasta* documenta alcance sobre oraciones (59% de los casos) (28a), sintagmas preposicionales (28%) (27b), sintagmas adjetivales (2%) (27c), sintagmas adverbiales (2%) (27d) y predicados (2%) (27e):

- (27) a. Se pochan y hasta [huelen mal]. (CORPES XXI, escrito)
- b. que tenía que cambiar de vida, de casa, de moto y hasta [de cabeza] (CORPES XXI, escrito)
- c. O como Gonzalo el policía, tan educado, hasta [tan cortado] (CORPES XXI, escrito)
- d. Me siento hasta [mal] y con la tripa revuelta. (CORPES XXI, escrito)²¹
- e. la situación empezaba a superarla y andaba ya hasta [tradicionando sus convicciones más profundas sobre una relación]. (CORPES XXI, escrito)

También se documentan ejemplos en la muestra escrita en los que se satisface la presuposición de alternativas sin que haya mención de alternativas en el discurso inmediato. La partícula muestra alcance sobre oraciones (55%) (28a), sintagmas preposicionales (39%) (28b), predicados (3%) (28c) y subordinadas (3%) (28d)²²:

- (28) a. No sabes tú bien la que quieren montar, y podrás imaginarte que lo conozco de primera mano. Vamos, que hasta [he visto el proyecto, incluidos los planos y el vídeo de la pirámide][...]. (CORPES XXI, escrito)

21. En este ejemplo, cabe interpretar también que «hasta» tiene alcance sobre el SP, y que, por tanto, el segundo miembro de la coordinación no es realmente una alternativa.

22. A diferencia de los otros ejemplos de (28), la subordinada de (28d) tiene una posible alternativa codificada en el discurso inmediato en el cuantificador *todo*.

- b. en ese momento resultaba incómodo hasta [para él]. (CORPES XXI, escrito)
- c. Pues me ha explicado el infeliz que ha llegado hasta [a pensar en robar]. (CORPES XXI, escrito)
- d. La televisión lo coge todo, hasta [lo que no debe]. (CORPES XXI, escrito)

6. ANÁLISIS DE LOS RESULTADOS

Los resultados obtenidos en las secciones anteriores ponen de manifiesto que *hasta* focal es un activador de presuposición de alternativas. Esta presuposición se puede satisfacer al margen del tipo de segmento sobre el que tiene alcance la partícula, la mención o no de alternativas en el discurso inmediato y el tipo de medio, oral o escrito, del que proceden los ejemplos. Se observa, sin embargo, que no siempre se puede satisfacer la presuposición de alternativas asociada al foco y que los casos en los que esto sucede tienen lugar, sobre todo, cuando *hasta* se emplea en el discurso oral, tiene alcance sobre oraciones y no hay oraciones mencionadas en el discurso previo que sean comutables y, por tanto, interpretadas como alternativas del foco. En estos casos se observa que, aunque las condiciones del enunciado y del contexto discursivo no permiten satisfacer la presuposición de alternativas, el enunciado sigue siendo interpretable.

De este análisis se desprenden varias cuestiones que quedan abiertas y que merecen un estudio pormenorizado en investigaciones futuras. En primer lugar, cabe plantearse por qué el alcance de *hasta* sobre oraciones es el que, preferentemente, no permite satisfacer la presuposición de alternativas. Parece que la asociación del foco con sus alternativas es más sencilla cuando los elementos comparados están más restringidos semánticamente, bien por asociación (con otros elementos de una misma clase semántica o conceptual), bien por la semántica oracional en la que están insertos. Así, en una oración como *Resulta hasta aburrido* (CORPES XXI, oral), la semántica de *aburrido* nos permite establecer asociaciones con otros segmentos que son formalmente comutables y que, además, encajan en la semántica oracional (p. ej. *poco original*). En cambio, cuando el foco es toda la oración, los elementos que debemos comparar tienen una semántica más compleja (denotan tipos de situaciones) y están menos restringidos semánticamente (no hay palabras o constituyentes de los que dependan léxico-semánticamente, salvo las que se encuentren en otras oraciones del discurso previo). En estos casos, la mención de alternativas en el discurso inmediato será clave para satisfacer la presuposición de alternativas. De no haber esta mención, como sucede en el ejemplo *Tengo como un gran equipo / el año pasado hasta [hubo un momento que me parecía que*

*podía haber ganado la liga*_F (CORPES XXI, oral), las alternativas posibles para la oración focalizada son tantas y tan poco restringidas semánticamente, que no es posible satisfacer la presuposición de alternativas. En un ejemplo como este, solo cabe presuponer que las alternativas, de haberlas, además de ser oraciones, han de expresar una situación positiva, acorde a la afirmación previa de tener un gran equipo.

En segundo lugar, cabe plantearse por qué hay oraciones con *hasta* en las que no se satisface la presuposición de existencia de alternativas para el foco, pero siguen siendo semánticamente interpretables. Se ha de tener en cuenta que, desde el punto de vista teórico, las presuposiciones deben satisfacerse para que el contenido proposicional sea evaluable en términos de verdadero o falso. Una oración como *Compró hasta melón* será verdadera o falsa dependiendo de si la persona compró o no otras cosas.

En el caso que nos ocupa cabe plantearse dos posibles explicaciones: que la información presupuesta se acomode y, por tanto, se satisfaga en el contexto creado para tales efectos por el receptor del mensaje; o bien que la presuposición se cancele en favor de la implicatura de escalaridad del enunciado, que también activa *hasta*.

De acuerdo con van der Sandt (2012), la vía óptima para satisfacer la información presupuesta es encontrarla explicitada en el discurso previo. De no estarlo, las propias expresiones presuposicionales son indicadores de la clase de contexto en el que interpretar adecuadamente el enunciado. Si esto último no es posible, entonces el oyente construirá parte del contexto necesario para adecuar la información presuposicional. Este procedimiento es el que, de acuerdo con el autor, se denomina contextualización o acomodación (van der Sandt, 2012, p. 331).

De estar ante un caso de acomodación, cabe suponer que los ejemplos con *hasta* permiten presuponer alternatividad porque el hablante-receptor busca un contexto adecuado para interpretar la información presuposicional. Se ha de tener en cuenta que, de ser así, la presuposición de alternativas se acomoda en el plano de la evaluación, en la línea de lo que ya indicó Portolés (2011) (v. § 3): la alternancia se daría entre dos puntos de vista, el que considera la proposición con foco como menos probable frente al punto de vista que considera otras proposiciones como más probables.

También cabe plantearse la posibilidad de que, en los casos objeto de estudio, se haya producido una cancelación de la presuposición de alternativas en favor de la implicatura de escalaridad. *Hasta* es, como se indicó en las secciones 2 y 3, un activador de escalaridad, pues el enunciado en el que aparece induce a situar el segmento sobre el que tiene alcance la partícula en el extremo de una escala. Cuando el alcance es oracional, *hasta* intensifica el grado de improbabilidad del enunciado

y lo hace, por tanto, sobre un juicio o evaluación del hablante sobre el mensaje. Esa escala induce a presuponer la existencia de otros grados o medidas (más o menos probabilidad de otras situaciones), pero no de otras situaciones por las que se pueda comutar el segmento focalizado y de las que debamos presuponer su existencia. El hecho de que *hasta* tenga una semántica y una pragmática complejas puede explicar que, de no poder satisfacer la existencia de alternativas para el foco, el enunciado siga siendo interpretado al cancelarse la presuposición a favor de la implicatura.

7. CONCLUSIONES

En este artículo hemos revisado la presuposición de alternativas asociadas al foco con *hasta* en español europeo. En su uso como adverbio de foco inclusivo, escalar y evaluativo, *hasta* se ha relacionado en la bibliografía previa con la evocación, interpretación o implicación de alternativas, si bien no se ha examinado con detalle su comportamiento como activador presuposicional.

Con el propósito de analizar el comportamiento de *hasta* focal en un amplio abanico de ejemplos, en este artículo se han elaborado dos muestras, una oral y otra escrita, extraídas de CORPES XXI (versión 0.99). Estos ejemplos se han clasificado en dos fases. En la primera, los ejemplos orales y escritos se han tipificado teniendo en cuenta la posición de la partícula y el segmento sintáctico sobre el que tiene alcance; en la segunda, se han clasificado teniendo en cuenta si los ejemplos permiten o no satisfacer la presuposición de alternativas para el foco.

A partir del estudio de los ejemplos de la primera clasificación, se observa que *hasta* focal presenta las mismas pautas distribucionales en la muestra oral y en la escrita: la partícula precede al constituyente sobre el que tiene alcance, para el que se documenta, además, una amplia tipología sintáctica de segmentos (oraciones, sintagmas preposicionales, sintagmas nominales, sintagmas adjetivales, sintagmas adverbiales, sintagmas verbales, subordinadas y partes de perífrasis o quasi perífrasis).

En la segunda clasificación, se han analizado los ejemplos teniendo en cuenta las correlaciones entre la presuposición de alternativas para el foco y tres factores: el tipo de segmento sobre el que tiene alcance la partícula, la mención o no de alternativas en el discurso inmediato y el carácter planificado o no del texto (texto escrito frente a texto oral no planificado). En los casos en los que no se puede satisfacer la presuposición de alternativas para el foco, se observan tres tendencias interrelacionadas: *hasta* tiene mayoritariamente un alcance oracional, no hay alternativas mencionadas en el discurso inmediato y los ejemplos pertenecen a textos orales no planificados.

Aunque la explicación de estos resultados excede los propósitos de este artículo, hemos apuntado algunas de las posibles causas por las que, contra todo pronóstico, *hasta* no siempre permite satisfacer la presuposición de alternativas para el foco. Hemos señalado como primera causa de esta posibilidad el hecho de que *hasta* tiene, frente a otras partículas focales, una semántica y una pragmática complejas: además de aditiva o inclusiva, es escalar y evaluativa. Para los casos en los que no se satisface la presuposición de alternativas por las que conmutar el foco, se han ofrecido dos posibles explicaciones: la acomodación de la información presupuesta a un contexto evaluativo, creado *in extremis* por el receptor para poder interpretar el enunciado; o la cancelación de la información presuposicional a favor de la implicatura de escalaridad, de nuevo aplicada al plano de la evaluación.

De otra parte, hemos mostrado que estos usos de *hasta* son más frecuentes cuando la partícula tiene alcance oracional, no hay alternativas mencionadas en el discurso inmediato y este no está planificado, como sucede en los ejemplos de la muestra oral. El hecho de que el alcance oracional de *hasta* favorezca un uso no presuposicional de la partícula puede guardar relación con que las oraciones son tipos semánticos más complejos que los suboracionales, permiten asociaciones semánticas más abiertas y son, por tanto, difícilmente conmutables. Esta dificultad aumenta si no hay alternativas mencionadas en el discurso inmediato, algo esperable en el discurso no planificado.

Todo lo dicho explicaría, al menos provisionalmente, que *hasta* permita un uso no presuposicional aunque sea una partícula focal y que este uso se documente sobre todo en ejemplos procedentes de textos orales en los que la partícula tiene alcance sobre oraciones y no hay alternativas mencionadas en el discurso previo.

AGRADECIMIENTOS

La autora agradece los comentarios a los dos revisores anónimos del trabajo enviado a evaluación. Cualquier error que pueda aún contener este artículo es responsabilidad exclusiva de la autora.

FINANCIACIÓN

Proyecto *Partículas focales del español (I): caracterización y elaboración de una base de datos de referencia* (ref. 1380988), 2021-2022, financiado por la Universidad de Jaén dentro de la convocatoria correspondiente a 2020 de ayudas a proyectos de I+D+i en el marco del Programa Operativo FEDER Andalucía 2014-2020.

REFERENCIAS

- Albelda, M. (2014). Escalaridad y evaluación: rasgos caracterizadores de la intensificación pragmática. En E. Putska & S. Goldschmitt (Eds.), *Emotionen, Expressivität, Emphase* (pp. 79–94). Erich Schmidt Verlag.
- Briz, A., Pons, S. & Portolés, J. (coords.) (2008). *Diccionario de partículas discursivas del español* [en línea]. Disponible en <www.dpde.es> [última consulta 04/06/2022].
- Calhoun, S. (2009). What makes a word contrastive? Prosodic, semantic and pragmatic perspectives. En D. Barth-Weingarten, N. Dehé & A. Wichmann (Eds.), *Where prosody meets pragmatics* (pp. 53–78). Emerald.
- Dominicy, M. (1982). La evolución del español *hasta* en Hispanoamérica. *Anuario de Letras*, 20, 41–90.
- Feliú, E. & Pato, E. (2022). Descripción y comparación de *incluso* e *inclusive* en español actual. *Lengua y habla*, 26, 55–86.
- Fox, D. & Katzir, R. (2011). On the characterization of alternatives. *Natural Language Semantics*, 19, 87–107.
- Gotzner, N. (2017). *Alternative sets in language processing. How focus alternatives are represented in the mind*. Palgrave Macmillan Cham.
- Gotzner, N., Wartenburger, I. & Spalek, K. (2016). The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition*, 8, 59–95.
- Husband, E. M. & Ferreira, F. (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience*, 31(2), 217–235.
- Kim, Ch., Gumlogson, Ch., Tanenhaus, M. K. & Runner, J. T. (2015). Context-driven expectations about focus alternatives. *Cognition*, 139, 28–49.
- König, E. (1991). *The meaning of focus particles: A comparative perspective*. Routledge.
- Loureda, Ó., Cruz, A., Recio, I. & Rudka, M. (2021). *Comunicación, partículas discursivas y pragmática experimental*. Arco-Libros.
- Miyoshi, J.-N. (2015). La partícula española *hasta* y su función pragmática. *Lingüística*, 31(1), 7–23.
- Portolés, J. (2007). Escalas informativas aditivas: pruebas del español. *Spanish in Context*, 4(2), 135–157.
- Portolés, J. (2009). Alternativas convocadas por partículas discursivas. *Español Actual*, 92, 47–68.
- Portolés, J. (2011). Las partículas focales desde una perspectiva polifónica. En H. Aschenberg & Ó. Loureda (Eds.), *Marcadores del discurso: de la descripción a la definición* (pp. 51–76). Iberoamericana.
- RAE. Banco de datos (CORPES XXI) [en línea] (versión 0.99). *Corpus del Español del Siglo XXI (CORPES)*. Disponible en <<https://dle.rae.es>> [última consulta 06/04/2022].
- RAE. *Diccionario panhispánico de dudas* [en línea]. Disponible en <<https://dle.rae.es>> [última consulta 06/04/2022].
- RAE y ASALE (2025). *Nueva gramática de la lengua española*, vol. II. Madrid: Espasa.
- Rooth, M. E. (1985). *Association with focus*. Tesis doctoral. Amherst: University of Massachussets Amherst.

- Rooth, M. E. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75–116.
- Salazar, V. & Carretero, A. (2025). La cohesión en la Gramática Discursivo-Funcional: indagaciones a partir de la negación aditiva del español. *Círculo de lingüística aplicada a la comunicación*, 101, 257–270.
- Sánchez, C. (1999). Los cuantificadores: clases de cuantificadores y estructuras cuantificativas. En I. Bosque & V. Demonte (Dirs.), *Gramática descriptiva de la lengua española*, vol. 1 (pp. 1025–1128). Espasa.
- Torres, L. (2020). *The scalar focus operator hasta: An experimental study on processing costs in Spanish*. Tesis doctoral. Heidelberg: Universität Heidelberg.
- Van der Sandt, R. (2012). Presupposition and accommodation in discourse. En K. Allan & K. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics* (pp. 330–350). Cambridge University Press.



ESTUDIO PILOTO DE LOS FAMILECTOS EN ESPAÑOL: OPTIMIZACIÓN METODOLÓGICA E IMPLICACIONES PARA EL ANÁLISIS DE LA VARIACIÓN SOCIOLINGÜÍSTICA

María ALCEDO-ORTIZ¹
Universidad de Sevilla

Bárbara MARQUETA-GRACIA²
Universidad de Zaragoza

Antonio BENÍTEZ-BURRACO³
Universidad de Sevilla

Resumen

El *familecto* puede definirse como la variedad lingüística asociada a las unidades familiares, aquellos grupos de personas convivientes desde hace tiempo con las que se mantiene una relación de parentesco y que comparten un horizonte de experiencias común. Sin embargo, este término no ha sido bien acogido por la sociolingüística española y supone un área de investigación que plantea muchos interrogantes. En el presente estudio se han aprovechado las características tipológicas de las lenguas esotéricas (aquellas empleadas por comunidades pequeñas de hablantes) para ofrecer una caracterización preliminar del familecto en español. Las hipótesis formuladas al respecto se han cotejado con la reunión y análisis de un corpus piloto compuesto por conversaciones mantenidas por familias de dos localidades del área metropolitana de Sevilla (Gines y Dos Hermanas). Los resultados se han comparado además con las características del español coloquial, llegándose a la conclusión de que el familecto podría considerarse un sociolecto diferenciado.

Palabras clave: familecto; español coloquial; sociolingüística; tipología lingüística; variedades del español

-
1. malcedo@us.es;  <https://orcid.org/0009-0000-6743-9403>
 2. marqueta@unizar.es;  <https://orcid.org/0000-0002-2199-1689>
 3. abenitez8@us.es;  <https://orcid.org/0000-0003-4574-5666>

PILOT STUDY ON SPANISH FAMILCTS: METHODOLOGICAL OPTIMIZATION AND IMPLICATIONS FOR SOCIOLINGUISTICS

Abstract

The *familect* can be defined as the linguistic variety associated with family units, namely, groups of kin people living together. However, this term has not been accepted by Spanish sociolinguistics and represents a controversial, but interesting area of research. In this paper, the typological characteristics of esoteric languages (those used by small communities of speakers forming dense networks) have been used to sketch a preliminary characterization of the familect in Spanish. The resulting hypotheses have been checked via the collection and analysis of an original pilot corpus comprising daily conversations from 5 families living in two towns from the metropolitan area of Seville (Gines and Dos Hermanas). Results have been further compared with colloquial Spanish. We conclude that the familect should be regarded as a differentiated sociolect.

Keywords: familect; colloquial Spanish; Sociolinguistics; linguistic typology; Spanish varieties

RECIBIDO: 28/02/2025

APROBADO: 20/09/2025

1. INTRODUCCIÓN

En esta investigación se busca aplicar lo que la tipología sociolingüística ha indagado con respecto a las relaciones entre diversidad estructural y sociopolítica en las lenguas del mundo (en particular, sobre la naturaleza de las lenguas exóticas y esotéricas) al estudio y la interpretación de la variación intralingüística en una misma lengua. Para ello, se analizará y discutirá el caso de los familectos, esto es, las variedades usadas para la comunicación entre miembros de la misma familia. La discusión girará en torno a si los familectos pueden equipararse a las lenguas habladas por grupos humanos cerrados (esto es, a las lenguas esotéricas, siguiendo a Wray y Grace, 2007; y en particular, a Benítez-Burraco y Felú Arquiola, 2024), y si son variedades lingüísticas con características distintivas. Para ello, se ha reunido un corpus de conversaciones familiares, que se han analizado a diferentes niveles: fónico, morfológico, léxico y sintáctico⁴.

4. En la línea de los resultados obtenidos en los estudios de tipología sociopolítica (Chen et al., 2024) nuestro interés preliminar se ha centrado en el análisis de las estrategias lingüísticas de los planos mencionados y no tanto en los fenómenos de la enunciación, de la estructura informativa y de la interacción en el discurso, propios del nivel pragmático y con el que completaríamos el análisis pragmagramático de los familectos. Estas estrategias las dejamos para un futuro trabajo en el que se examinen minuciosamente los aspectos pragmático-discursivos.

El *familecto* (Wertheimer, 1973; Søndergaard, 1991⁵), también conocido en la literatura lingüística moderna como *lenguaje familiar*, *comunicación familiar cotidiana*, *fraseología doméstica*, *dialecto familiar*, *nociónes familiares*, *comunicación familiar*, *lenguaje familiar*, *léxico familiar* u *oikolecto*⁶, denota la variedad lingüística, en términos formales y discursivos, empleada en la unidad familiar. Presumiblemente, fue Wertheimer el primero en describir el fenómeno:

A familect is, as I mentioned before, a kind of microdialect. Most families have one; failure to use it in family settings implies that something is wrong somewhere, that the speaker is upset, or the like. And, of course, such a lect would never be used with someone outside the family. The idiosyncratic lexical items and usages may become so automatic that it is hard even to think of them in other contexts; they sound odd to outsiders but of course seem perfectly natural inside the family. (1973, p. 8)

La sociolingüística tradicionalmente ha considerado la comunicación entre familiares como parte del registro coloquial, al mismo nivel que las conversaciones con amigos, conocidos, allegados, o parientes no convivientes. Es indudable que son manifestaciones que presentan rasgos propios de este registro, como reflejaremos a continuación, pero dada la particularidad del grupo social en el que se adscribe, interesa identificar si en la conversación entre familiares se generan particularidades lingüísticas propias y diferentes a las de otros sociolectos. Nosotros, en cambio, trataremos de demostrar que la familia es un grupo social con características suficientes como para generar particularidades lingüísticas propias (§2.2).

A través de la comunicación, los seres humanos establecen, mantienen y disuelven sus relaciones socioafectivas, económicas, políticas y culturales (Gallego Uribe, 2006). En el caso de las familias, estas se centran en el mantenimiento de la jerarquía entre sus miembros, la organización del trabajo y de las actividades del hogar, el mantenimiento de los vínculos interpersonales entre sus integrantes y la puesta en común de ideas y de vivencias compartidas (cf. Gottman y Krottoff, 1989; Levenson y Gottman, 1985; Huston y Vangelisti, 1991; Vangelisti, 2004). Con ello, y como ha señalado Gordon (2009), se refuerzan dos funciones vitales para la pervivencia del grupo: la función identitaria, que favorece la cohesión y

5. Søndergaard emplea el término en contraposición al idiolecto y sociolecto para analizar el comportamiento plurilingüe en el lenguaje hablado de una familia políglota.

6. «...familect, known in the modern linguistic literature as 'family language', 'family everyday communication', 'domestic phraseology', 'family dialect', 'family notions', 'family communication', 'family language', 'family lexicon', 'oikolect». (Osovská, 2019, p.124)

el sentido de pertenencia al núcleo familiar; y la función lúdica, que refuerza el afecto y la complicidad entre sus miembros.

El español es una lengua tipológicamente *exotérica* (Chen et al., 2024), pues su comunidad de hablantes es amplia y laxa, con muchos puntos de contacto entre personas, países y ramas técnicas, y cuenta con aprendientes en diferentes partes del mundo. Esta estructura social explicaría, según propuestas como las de Wray y Grace (2007), que su gramática sea sintácticamente compleja, con tendencia a la hipotaxis; su vocabulario sea amplio y variado, y posea una notable transparencia semántica; y que su grado de composicionalidad sea tan elevado. Sin embargo, es posible que los factores extralingüísticos que hacen de las lenguas códigos exotéricos (o esotéricos, que presentarían las características contrarias), favorezcan también la existencia de variedades esotéricas y exotéricas dentro de una misma lengua. Así, como sugieren Wray y Grace (2007), si el grupo de hablantes se reduce, las presiones culturales aumentan y lo hacen también los factores identitarios, generaría características gramaticales propias de la esoteridad. Si esta hipótesis se confirma, supondría un punto de inflexión en la teoría variacionista al plantear un nuevo presupuesto teórico para el análisis de las variedades diastráticas del español.

Expuesto este marco, surgen tres cuestiones que sentarán la base de los objetivos de esta investigación: a) ¿la reducción del núcleo social que implica la unidad familiar conlleva una variación de tipo sociolingüístico?; b) ¿esas particularidades se pueden interpretar en términos de una comunicación esotérica y asemejarse a las que distinguen a las lenguas de esta clase?; c) ¿se podría defender que el familecto es una variedad sociolingüística diferenciada? En consonancia con estos interrogantes, los objetivos de esta investigación se resumen en cuatro:

1. Reunir un corpus de conversaciones familiares que permita comprobar estas hipótesis.
2. Caracterizar de forma unitaria la variedad lingüística usada en las interacciones familiares.
3. Discutir si dicha variedad (el familecto) presenta diferencias con respecto al español coloquial lo suficientemente distintivas como para defender que constituye una variedad lingüística propia (sociolecto).
4. Comprobar si los rasgos encontrados se corresponden con los que se asocian a las lenguas esotéricas.

2. MARCO TEÓRICO

Resulta pertinente examinar cuáles son las características lingüísticas que típicamente se han asociado al estilo coloquial del español, para poder luego discriminar entre los rasgos propios del estilo informal y conversacional, y los que serían idiosincrásicos de los familectos. Es también útil analizar sucintamente los pormenores de su conceptualización y el marco teórico en el que se inserta tal español coloquial, y conocer, en particular, qué variables contextuales y sociales se aplican habitualmente para la delimitación de este estilo, siempre con el objetivo de comparar todo lo anterior con la familia. Por otro lado, también resulta adecuado examinar el estado de la investigación sobre las lenguas exóticas y esotéricas, para entender en qué sentido el familecto puede considerarse una variedad del segundo tipo.

2.1. *El estilo coloquial en español: características lingüísticas y debate conceptual*

El español coloquial ha sido caracterizado lingüísticamente en gran detalle por autores como Briz (1996, 1998), Narbona, Sánchez y Lázaro Carreter (1993), Cortés Rodríguez (1991, 2002) o Gutiérrez Ordóñez (1997). Gaviño (2023), que recoge estos trabajos, resulta idóneo para nuestro propósito, porque atiende a los rasgos por planos o niveles de la lengua: fónico, morfológico y léxico, y sintáctico, que es precisamente el enfoque que pretendemos utilizar para analizar los resultados del corpus. De esta manera, las características lingüísticas del español coloquial se resumen en:

NIVEL FÓNICO

- a) Alargamientos de sonidos, que pueden tener diversas funciones (o ser un rasgo dialectal, sin función pragmática): recurso de apoyo para organizar la información; mantenimiento del turno de palabra; modificación de la fuerza ilocutiva; función atenuadora, etc.
- b) Metátesis o cambios de sonidos: de nuevo como rasgo dialectal o idiolectal; pero pueden obedecer también a causas discursivas, como la rapidez de locución; o involuntario.
- c) Debilitamientos y pérdidas: aféresis (p.e. *Ángela* > *Gela*); apócope; síncopa.
- d) Adiciones de sonidos: epéntesis; prótesis; paragoge.
- e) Elementos suprasegmentales y no verbales con funciones comunicativas diversas: énfasis, tipos de emoción, etc.

NIVEL MORFOLÓGICO Y LÉXICO

- a) Empleo de *palabras comodín* u *ómnibus*: proverbios (p.e. *hacer, tener, haber, poner, decir, echar, dar*), prosustantivos (*cosa, rollo, chisme, tío*), proadjetivos (*estupendo, chungo, flipante, fuerte, guapo, guay*), o proadverbios (*super, bien, muy*).
- b) Preferencias léxicas. Gaviño (2023) enumera algunas: *sitio* por *lugar*; *mejor* por *preferible*; *así* por *de este modo*; *casi* por *apenas*; *ropa* por *prenda*; *pero* por *sin embargo*; *partir* por *romper*, *raro* por *extraño*, *a lo mejor* por *quizás*; *tener ganas de* o *querer* por *desear*, etc.
- c) Redundancias léxicas. Aunque también aparecen en el estilo formal, son más comunes en el informal. Pueden tener la función pragmática de fortalecer un contenido o aumentar su expresividad.
- d) Lenguaje formulario: fraseología; expresiones fijas; locuciones y modismos de carácter idiomático.
- e) Procedimientos de creación léxica: derivación, composición, acortamiento léxico, usos metafóricos o metonímicos.
- f) Léxico del insulto. El léxico del insulto y toda su fraseología merece una explicación más exhaustiva. Los hablantes tienden a emplear de manera asidua palabras o expresiones marcadas como malfacientes dentro de la norma de la lengua. Las causas sociolingüísticas subyacentes son muy variadas: necesidad de herir verbalmente a otros, necesidad de relajar los momentos de tensión en las conversaciones, o alabanza del otro (valor encomiástico). Los insultos son también recursos expresivos que usamos como reguladores del acercamiento social (una función característica entre los jóvenes) o como muestras directas de un alto grado de familiaridad. En tales casos, pierden su valor originario denigrativo.

NIVEL SINTÁCTICO

- a) Sintaxis parcelada: preferencia por construcciones coordinadas y yuxtapuestas.
- b) Variedad en el orden de constituyentes: topicalizaciones y focalizaciones, y progresión remática y temática; construcciones sintácticas que distan de la norma.
- c) Elisiones de constituyentes. Los cambios en el orden sintáctico se compensan por medio de estrategias de cohesión.

- d) Anacolutos o cambios repentinos, conscientes e inconscientes, en la construcción del enunciado. Rompen las uniones sintácticas entre los elementos.
- e) Enunciados suspendidos⁷.
- f) Variedad de funciones de los nexos coordinantes y subordinantes.

Tradicionalmente se ha considerado que la sintaxis coloquial responde a una forma de expresión reducida, pobre o poco refinada, posiblemente por las formulaciones de Beinhauer (1968). Como afirma Narbona (2012, 2018), es equivocado ligar la espontaneidad a la irregularidad lingüística. De hecho, se acepta en la actualidad que «no hay razones lingüísticas fundadas para achacar una falta de complejidad y estructuración en la estructura sintáctica de los enunciados coloquiales, ya que la lengua coloquial es quizás desde el punto de vista gramatical la más compleja de todas» (Gaviño, 2023, p.95).

Los estudios sobre coloquialidad y oralidad se insertan en la teoría de las variedades de la lengua o variedades de uso de la lengua, en concreto en el nivel diafásico, donde se analizan las relaciones entre las elecciones lingüísticas y la situación extralingüística. La propuesta inicial de Halliday, McIntosh y Strevens (perfeccionada por Gregory y Carroll, 1978), incluía, tal como señala López Serena (2007), cuatro aspectos o dimensiones para delimitar los diferentes registros y estilos: *campo* del discurso (tema o tópicos del discurso), *modo* del discurso (espontáneo, preparado, memorizado, etc.), *tono funcional* del discurso (ligado a las modalidades discursivas) y *tono personal* del discurso (grados de formalidad), generando dos tipos fundamentales de estilos: formal e informal. Los estudios iniciales de la coloquialidad en lengua española ligaron la modalidad coloquial al estilo informal (cf. Seco, 1972; Herrero, 1990; Briz, 1996, 1998). Briz, en particular, considera que el elemento determinante, y lo que confiere identidad a este estilo, es la particularidad de la situación comunicativa, por lo que los rasgos situacionales (o coloquializadores) serían los que motivan las características del estilo coloquial (Figura 1).

7. Briz define los enunciados suspendidos como «entendidos como tales en cuanto presentan un valor ilocutivo completo o, más exactamente, completado en el contexto por el oyente, a diferencia de los simples cortes en lo comunicado, de los abandonos de una construcción para empezar con otra, reinicios, etc.» (1998, pág. 86), definición en la que resalta lo ilocucionario como criterio identificador. Por otro lado, Pérez Béjar define y clasifica los enunciados suspendidos según criterios sintácticos: «corresponde a estructuras comparativas, consecutivas, condicionales, causales o respectuales truncadas. Falta el segundo miembro, pero pervive el nexo o parte de él. Este, junto con la entonación ascendente, constituye la marca de la construcción». (2018, pág. 42). Será el criterio de Pérez Béjar el que sigamos en el análisis.

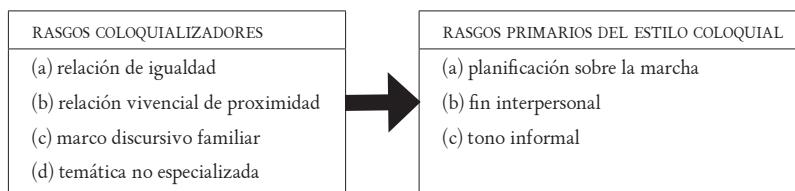


Figura 1. Rasgos primarios del estilo coloquial de acuerdo con los rasgos situacionales que lo caracterizan propuestos por Briz (cf. 2010, p. 131). La *relación de igualdad* (a) se refiere a la relación interpersonal entre los interlocutores, ya sea social (determinada por el estrato sociocultural) o funcional (el papel o rol que cada uno desempeña en la comunicación). La *relación vivencial de proximidad* (b) atiende al saber y al número de experiencias compartidas. El *marco discursivo familiar* (c) está delimitado por la relación de los participantes con el espacio o lugar donde se mantiene la comunicación. La *temática no especializada* (d) se vincula a los temas cotidianos.

Las aportaciones de los romanistas alemanes Koch y Oesterreicher también resultan útiles para delimitar la coloquialidad. Su modelo (véase Koch y Oesterreicher 1990/2007) defiende que los registros lingüísticos se sitúan a lo largo de un *continuum*, polarizado entre la inmediatez (la oralidad) y la distancia comunicativa (la escrituralidad), y que se caracteriza por un haz de valores paramétricos que describen las diferentes condiciones comunicativas⁸. El polo de la distancia se caracteriza por discursos planificados, formales y predominantemente escritos; mientras que en el polo de la inmediatez se sitúan las manifestaciones orales, espontáneas y altamente informales, en las que los interlocutores comparten un contexto situacional y emocional estrecho, que permite un mayor grado de cooperación, una comunicación menos mediada y un menor grado de planificación.

8. Koch y Oesterreicher plantean hasta once condiciones comunicativas (Figura 2). Sin embargo, no conciben un inventario cerrado, sino que invitan a la proposición (y, por ello, añaden un etcétera al final de la enumeración). La condición *grado de publicidad* (a) denota el carácter público o privado de la interacción, para el que resulta relevante el número de interlocutores y la existencia o no de público, así como sus dimensiones. Por *grado de familiaridad* (b) hay que entender el nivel de confianza entre los interlocutores. El *grado de afectividad* (c), el nivel de implicación emocional con respecto al interlocutor. El *grado de expresividad* (d) se relaciona con el nivel de implicación emocional con respecto a lo que se comunica. C) y d) son, por tanto, opciones de un mismo componente, a saber, el nivel de implicación emocional, que se regiría, así, por esos dos elementos. El *grado de anclaje comunicativo* (e) tiene que ver con la situación y con la acción comunicativas. La condición *tipo de referencialidad* (f) alude a la posibilidad de referencialización desde el presente comunicativo (el aquí y ahora del hablante). La *distancia* (g) se entiende como medida física. Finalmente, *grado de fijación temática* (k) se vincula con el modo en que un discurso o conversación mantiene el enfoque sobre un determinado tema. Si la fijación es fuerte, los textos o discursos son formales y estructurados. Si la fijación es débil, los discursos o conversaciones se consideran informales, con lo que el enfoque cambia constantemente.

Son estos parámetros de la situación comunicativa los que motivan las diferentes estrategias de verbalización (Figura 2).

CONDICIONES COMUNICATIVAS	ESTRATEGIAS DE VERBALIZACIÓN
<ul style="list-style-type: none"> (a) grado de publicidad (b) grado de familiaridad (c) grado de afectividad (d) grado de expresividad (e) grado de anclaje comunicativo (f) tipo de referencialización (g) distancia entre interlocutores (h) grado de cooperación (i) carácter dialógico o monológico (j) grado de espontaneidad (k) grado de fijación temática etc. 	<ul style="list-style-type: none"> (a) grado de contextualización lingüística o extralingüística (b) grado de planificación (c) carácter provisional o definitivo (d) sintaxis agregativa o integrativa etc.

Figura 2. Condiciones comunicativas y estrategias de verbalización posibles en la distancia y en la inmediatez comunicativas según Koch y Oesterreicher (1990/2007, p. 182)

Tanto Briz como Koch y Oesterreicher coinciden en poner de relieve como rasgos definitorios de la coloquialidad la pertinencia de la relación entre los interlocutores, la relación de estos con la situación o espacio comunicativo, la libertad temática y la ausencia de planificación o espontaneidad de la comunicación. Destacamos los conceptos de *relación vivencial de proximidad* y de *marco discursivo familiar* de Briz porque, creemos, pueden vincularse expresamente con algunas de las particularidades sociológicas que atribuimos a la identidad sociolingüística familiar: el amplio número de experiencias compartidas y la relación directa que mantienen los participantes con el espacio y el lugar de la comunicación, especialmente endógeno en el caso familiar. Queda por discutir si en la conversación familiar priman temas no especializados y de la vida cotidiana o, por el contrario, se registran temas variados y de distintos grados de especialización. Entre las condiciones comunicativas del modelo de los investigadores alemanes, el familecto se caracterizaría por un grado elevado de muchas de las condiciones comunicativas propias de la inmediatez comunicativa, como un amplio grado de familiaridad entre los interlocutores, un elevado grado de implicación emocional interpersonal (afectividad), el máximo anclaje conversacional en la situación, además de un campo relacional limitado al hogar, inmediatez física de los interlocutores, dialogicidad y espontaneidad y

alto grado de cooperación. A diferencia del modelo de Val.Es.Co, que propone una temática no especializada, el modelo variacional de Koch y Oesterreicher no cierra la puerta a la especialización temática, sino que evalúa el grado de mantenimiento de los temas de conversación. Tras el análisis de los datos, discutiremos si se puede defender el familecto como un sociolecto diferenciado y si los parámetros propuestos por estos dos modelos situacionales reflejan adecuadamente la particularidad del fenómeno.

2.2. *Familecto: ¿una forma de comunicación diferenciada?*

En el marco de los estudios multiculturales y plurilingües, Sønsgaard (1991) introduce el concepto de *familect* (familecto), con el que denota a la variedad lingüística compartida por un grupo de hablantes que son miembros de una misma familia. Sønsgaard emplea el concepto para analizar el multilingüismo, no desde una perspectiva idiolectal o sociolectal, como se venía haciendo entonces, sino como parte de un entramado más íntimo, grupal y, sobre todo, distintivo, en el que todos los miembros de un grupo reducido se ven lingüísticamente implicados. Este código lingüístico crea identidad entre los miembros del grupo y los hace reconocibles respecto a los hablantes externos. Es a raíz de trabajos como el de Sønsgaard del que parte el interés por la comunicación íntima o familiar.

En años posteriores, la sociología se ha interesado por la comunicación familiar, teniendo como objetivo averiguar qué características sociológicas distinguen a la familia de otros tipos de relaciones sociales (Vangelisti, 2004; Galvin et al., 2004; Gallego Uribe, 2006) y descubriendo que la comunicación refleja los vínculos interpersonales entre los integrantes (Gottman y Kroftoff, 1989; Levenson y Gottman, 1985)⁹, revela la clase y calidad de las relaciones familiares (Huston y Vangelisti, 1991) y devuelve patrones lingüísticos que simbolizan estrategias intrafamiliares (Canary y Stafford, 1992).

Nosotros, como lingüistas, debemos continuar esta línea de investigación y explorar la variación intralingüística puede ser una oportunidad para hacerlo. Para ayudarnos a trazar una propuesta de caracterización lingüística, podemos servirnos de trabajos previos que hayan tratado de determinar los rasgos lingüísticos, comunicativos y sociológicos asociados a las interacciones verbales en el seno de las familias. Clancy (2016) señala como uno de los rasgos distintivos de este tipo

9. Estos trabajos son anteriores al de Sønsgaard (1991), pero debemos incluirlos por la amplia contribución que realizan al estudio de las relaciones de matrimonio.

de interacciones la idiomática, que se correlaciona con otros rasgos también singulares como son la intimidad y la complicidad entre los participantes. En relación con esta elevada idiomática, Gordon (2009) expone que cada familia tiene su propio repertorio léxico, caracterizado por el juego lingüístico y por la presencia de la imitación del lenguaje infantil. Esto podría ocurrir, como reflexionan Benítez Burraco y Felíu-Arquiola (2024), por la influencia del *habla infantil*, que es la variedad lingüística que desarrollan los niños durante el proceso de adquisición de la lengua materna y que se caracteriza por todo tipo de simplificaciones fonológicas, morfológicas, semánticas y sintácticas (Lust, 2006; Roberts, 2013). Pero otro factor podría ser la influencia de los estilos discursivos que adoptan los adultos que interactúan durante mucho tiempo con niños pequeños, conocido en la literatura como *maternés* o *habla del cuidador* (Kempe et al., 2007; Schick et al., 2022). Finalmente, no puede olvidarse la infantilización inherente que conllevan, en muchos casos, las relaciones afectivas. Gordon (2009) también destaca aspectos como la intertextualidad y la ventrilocuización, que provocarían la repetición total o parcial de fragmentos discursivos de otros miembros de la familia con tendencia a la imitación de su *modus loquendi*.

Además del juego lingüístico, el repertorio léxico propio de cada familia estaría también marcado por la función identitaria, que vendría alimentada por el amplio número de experiencias y referencias compartidas, y por un elemento que sí es exclusivo de este grupo social, a saber, la consanguinidad, que conlleva una serie de implicaciones biológicas y legales, además de sociales y culturales (Arés Muzio, 2003). Esta función identitaria confiere un valor idiosincrásico a la comunicación, generando significados y formaciones léxicas que solo son conocidos por los miembros de la familia (y que se reservan la opción de compartir con los no pertenecientes al núcleo familiar). Osovská (2019), quien ha caracterizado el léxico de las familias alemanas, señala que se compone de apelativos particulares, palabras jergales o argóticas y términos de nueva creación, aunque generados mediante procedimientos de formación de palabras ya establecidos en la lengua.

En resumen, podríamos decir que en la variedad hablada familiar priman las funciones identitaria y lúdica, de modo que en los familectos se integrarían las estrategias verbales que las familias ponen en práctica para satisfacer tales funciones comunicativas. Todas estas ideas se sintetizan en la Figura 3.

FUNCIONES COMUNICATIVAS	GORDON (2009)	CLANCY (2016)	Osovská (2019)
Función identitaria	(a) carácter ritualizado (b) intertextualidad y ventriloquización	(e) grado de idiosincrasia (f) grado de intimidad (g) grado de complicidad	(h) apelativos particulares, palabras jergales o argoticas
Función lúdica	(c) juego lingüístico (d) imitación del habla infantil	x	(i) abundantes indicios de juego léxico

Figura 3. Características del discurso o habla familiar expuestas por Gordon (2009), Clancy (2016) y Osovská (2019) en relación con las funciones comunicativas que se le atribuyen.

2.3. Familecto: ¿una forma de comunicación esotérica?

Finalmente, y en la línea también de lo apuntado en el apartado introductorio, las características discutidas en los dos subapartados anteriores invitan a trazar un paralelismo entre las manifestaciones lingüísticas familiares y las interacciones esotéricas, y en particular, las lenguas esotéricas o lenguas de tipo S (en la terminología de Chen et al., 2024)¹⁰. Variedades familiares y lenguas esotéricas coinciden en el elevado grado de referencialidad extralingüística y de dependencia contextual, que se relaciona con el mayor indicio de implicaturas. Benítez Burraco y Felú Arquiola (2024) han hecho una propuesta de caracterización de los familectos del español a partir de las propiedades distintivas de las lenguas esotéricas o lenguas de tipo S. Cabe recordar que esta propuesta de caracterización se centra en aspectos lingüísticos, los mismos que analizaremos en el presente estudio, en sintonía con los resultados que arrojan los recientes estudios de tipología sociopolítica (cf. Chen et al., 2024). Los familectos en nuestra lengua tendrían, así, las siguientes características:

NIVEL FÓNICO

- Modificaciones secundarias de sonidos habituales, como la palatalización (por ejemplo, *chí* por *si*).
- Alteraciones prosódicas, como cambios en el patrón acentual (por ejemplo, acortamientos que produzcan formas bisílabas de acentuación llana: *Miguel* > *Migue*).

10. Fue Givón (1979) quien comentó que cuanto más reducido es el grupo de hablantes y más vínculos genéticos unen a sus miembros, menos información transmiten y más elementos lingüísticos eliden por compartir propósitos y actividades comunes. Añade este autor que dicha particularidad idiomática podría estar en el origen de las lenguas prehistóricas y que el crecimiento de las sociedades provocaría su desnaturalización hacia sistemas más flexibles e idiomáticamente más abiertos (cf. Kay, 1977).

- c) Alteraciones de la fonotaxis habitual, en particular, el incremento de la complejidad silábica (como en *sip* por *si*) o del número de sílabas mediante ampliación silábica o reduplicación (por ejemplo, *vayavaya* por *vaya* con valor enfático).
- d) Juegos fónicos en general: aliteraciones, rimas, paronomasias, etc. (por ejemplo, en expresiones como *hola, caracola* o *de nada, monada*).

NIVEL MORFOLÓGICO

- a) Procesos derivativos como el empleo frecuente de sufijos apreciativos y de prefijos intensivos (por ejemplo, *fotaza* o *me requetechifla*); sufijos propios del lenguaje juvenil (por ejemplo, *careto*), propios de ámbitos especializados (por ejemplo, *mamitis*) o de creación propia; o sufijos propios de la lengua general con bases inesperadas (por ejemplo, *peinaje* para designar la acción de peinar o *dormición* para la acción de dormir).
- b) Compuestos de nueva creación (por ejemplo, *posamóvil*, *cagarrisas*).
- c) Empleo de formas irregulares (por ejemplo, extensión del sufijo culto *-érrimo* a casos donde no es esperable, como en *malérrimo*).
- d) Usos no habituales del género y el número gramaticales, como el empleo del masculino despectivo o desmerecedor (por ejemplo, *coso* o *palabro*).

NIVEL SINTÁCTICO

- a) Orden de constituyentes variado, distinto de SVO, por razones que no sean las puramente informativas (por ejemplo, *Viene María tarde*).
- b) Formación de interrogativas *in situ* (por ejemplo, *¿Has comido qué?*).
- c) Construcciones con función antipasiva (por ejemplo, *Juan aprovecha tu propuesta > Juan se aprovecha de tu propuesta*).
- d) Cambios de diátesis sin marca (por ejemplo, *Los garbanzos congelan bien*).
- e) Predominio de las oraciones simples sobre las complejas (por ejemplo, *Mi amigo vino ayer. Se llama Juan*, frente a *Mi amigo, que se llama Juan, vino ayer*).
- f) Preferencia por la yuxtaposición (por ejemplo, *Voy a abrir la ventana, hace calor* frente a *Voy a abrir la ventana porque hace calor*).

NIVEL LÉXICO

- a) Combinaciones léxicas poco frecuentes (por ejemplo, alteración de asociaciones léxicas habituales, como *rebanada de queso* o *loncha de pan* en lugar de *rebanada de pan* o *loncha de queso*).
- b) Uso de modismos, especialmente frases hechas de creación propia y significado solo accesible para los miembros de la unidad familiar.

- c) Palabras de uso general que se emplean con un significado idiosincrásico (por ejemplo, *el huevo* para referirse a un coche).
- d) Presencia de marcas diásimas: coloquialismos, dialectalismos.
- e) Presencia de apodos e hipocorísticos (por ejemplo, *el tete* para el hermano, *yayo* o *yaya* para los abuelos).

3. METODOLOGÍA Y CORPUS

3.1. *Metodología*

Estudiar la comunicación familiar conlleva varios desafíos, como el hecho de que los sistemas familiares no son fijos, sino que evolucionan a través del tiempo (es decir, cuentan con etapas de desarrollo); o que sus miembros son de diferente nivel sociocultural y mantienen relaciones diferentes entre ellos (es decir, cuentan con diferentes diádicas constitutivas). Todo ello, sumado a los contextos tan privados en los que ocurren las interacciones, dificulta la obtención y la interpretación de los datos. Para evitar la paradoja del observador y tomar muestras aún más naturales, se ha encargado a un miembro de cada familia realizar las grabaciones, quedando a su criterio decidir cuándo comunicar al resto de integrantes la participación en el estudio. Solo se han aceptado a las familias que han manifestado su consentimiento expreso de participar en el estudio.

La recogida de muestras se ha llevado a cabo mediante un dispositivo digital, generalmente un teléfono móvil, que permite grabar audios durante un periodo de tiempo prolongado y que fue colocado de modo que no restringiera la entrada de audio. Se ha grabado durante periodos prolongados en diferentes momentos del día (en general, cuando el volumen de temas de conversación era mayor) y preferiblemente a las mismas personas. Los intervenientes son quienes comparten un parentesco cercano y conviven desde hace tiempo. No se han grabado conversaciones en las que participaran niños menores de 8-10 años para evitar las interferencias con el lenguaje infantil. De los intervenientes se han solicitado una serie de datos de interés sociolingüístico, así como una relación de términos que consideraran idiosincrásicos de su unidad familiar. En caso de padres divorciados, se ha escogido un solo núcleo familiar.

La transcripción de los datos se ha realizado en dos pasos. Primero, se ha hecho una transcripción automática mediante un programa de reconocimiento de voz (<https://github.com/mario-casado-m/whisper-webmaus>). Seguidamente,

la transcripción ha sido revisada ortográfica y gramaticalmente para corregir los previsibles errores cometidos por el programa. A la transcripción se han añadido comentarios sobre aspectos paralingüísticos y contextuales de interés, con objeto de entender mejor el sentido de las interacciones. Conviene aclarar que no se ha seguido un sistema de transcripción fonológico, pues los rasgos que nos interesaba comprobar son perfectamente representables con una transcripción ortográfica convencional, como las del grupo Val.Es.Co (Briz et al. 2002). El sistema seguido para anotar los elementos extralingüísticos más habituales puede consultarse en la Tabla 1.

Paréntesis ()	Señalan acciones que realizan los intervinientes y que ayudan a entender la comunicación.
Corchetes []	Reconstruyen enunciados difíciles de comprender por la presencia de ruido externo en la grabación.
Equis xxxx	Fragmentos incomprendibles en los audios
Negrita	Describen el contexto situacional.
Cursiva+paréntesis	Explican referentes concretos.
Comillas	Señalan el estilo directo o indirecto cuando el número de enunciadores.
Asteriscos	Especifican algún hecho puntual y significativo del discurso.

Tabla 1. Sistema de anotación utilizado para comentar aspectos extralingüísticos en el corpus.

3.2. *Corpus*

Debido al carácter de estudio preliminar o piloto, el corpus se ha conformado a partir de solo cinco familias, pertenecientes al área metropolitana de Sevilla, a las que se pidió que grabaran sus conversaciones durante cinco momentos diferentes del día y por períodos de aproximadamente veinte minutos (por grabación), procurando recoger, en la medida de lo posible, conversaciones completas.

La primera familia se compone de tres miembros: el padre (57 años), la madre (60 años) y una hija de 23 años. La madre y la hija poseen estudios universitarios, mientras que el padre posee estudios de grado medio. Su situación económica es media-alta y viven en la localidad de Gines (Sevilla).

La segunda familia se compone de cuatro miembros, aunque solo tres de ellos aparecen en las muestras, ya que el hermano mayor se emancipó hace un año. Estos son: el padre (58 años), la madre (55 años) y el hijo menor (23 años). Todos poseen estudios universitarios y su situación económica es media-alta. Viven en la localidad de Dos Hermanas (Sevilla).

En la tercera familia los padres están divorciados, por lo que solo se han recopilado grabaciones de uno de los núcleos familiares, el del padre, por elección del contacto. El núcleo familiar paterno se compone del padre (60 años), una hija (23 años) y un hijo (20) años. Tanto padre como hija tienen estudios universitarios, el hijo aún los cursa. Su situación económica es media-alta y viven en Gines (Sevilla).

La cuarta familia se compone de cuatro miembros: el padre (59 años), la madre (56 años), la hija mayor (23 años) y la hija menor (20 años). Todas poseen estudios universitarios, mientras que el padre cuenta con un grado medio. Tienen una situación económica medio-alta y viven en la localidad de Gines (Sevilla).

La quinta familia está formada por cuatro personas: el padre (57 años), la madre (55 años), el hijo mayor (28 años) y el hijo menor (26 años). El hijo mayor ya se ha emancipado, pero frecuenta el hogar familiar. Los dos hijos cuentan con estudios universitarios. El padre tiene estudios de grado medio y la madre, la formación básica obligatoria. Su nivel de renta es medio-alto y residen en Gines (Sevilla).

4. RESULTADOS Y ANÁLISIS

Al inicio del artículo, anticipábamos que el objetivo de análisis del corpus consiste en determinar si este presenta elementos que permitan caracterizar el sociolecto como una variedad lingüística distintiva y si, por otra parte, estos rasgos se corresponden con los que se asocian a las lenguas esotéricas.

Los resultados, como se ha venido advirtiendo, se presentarán divididos por planos o niveles de la lengua. Los ejemplos extraídos del corpus se señalarán de forma numerada y en cursiva. Al final de la cita se indicará la familia (F) a la que pertenece, numerada del 1 al 5 y la conversación de la que se ha extraído, también numerada del 1 al 5. Así, una cita extraída de la primera familia y de la tercera conversación se indicará como *F1-3*.

4.1. *Nivel fónico*

No hay muestras de variación diatópica ni diastrática en la pronunciación de las palabras, sino que obedecen a la variedad meridional y andaluza y, en concreto, a la sevillana. La ausencia de cambios en la pronunciación de vocales y consonantes, debilitamientos y pérdidas de sonidos, adiciones o metátesis fonémáticas que se alejen de la norma dialectal indica que las familias se mueven en un registro sociocultural medio-elevado, acorde con su nivel de estudios y economía familiar. El recurso de la repetición está muy extendido y parece emplearse con distintas

intenciones, desde la atenuación, hasta la intensificación, el énfasis o la aclaración. Encontramos un caso claro de asimilación contextual (3) y dos ejemplos de alteración de la fonotaxis habitual (1, 2), aunque parecen ser puntuales y consecuencia de la rapidez e instantaneidad del contexto enunciativo.

- (1) FAT[08:51.320 --> 08:56.320] Yo llevo los *ter-nedores*, llevo los platos y llevo los vasos. *F3-1*
- (2) DAU[05:44.240 --> 05:46.240] A ver, suena un poco entre el *cacahuate*. *F3-2*
- (3) DAU[08:10.320 --> 08:12.320] *Pon (pues)*, ponlo. *F3-2*

En el nivel fonológico, destacan dos elementos poco frecuentes en otros corpus orales y que no se han contemplado como rasgos del español coloquial, ni se han planteado en el protocolo de rasgos esotéricos. Uno de ellos es la presencia de una especie de sonido gutural agudo (marcado como **ruidos**) que realizan el padre y la hija de la familia 1 y que llama la atención por no ser ningún tipo de fonema del español, pero que, sin duda, lo emplean con un claro significado conversacional y quizás idiomático (4). El otro fenómeno afecta a la prosodia y consiste en la variación puntual de la entonación declarativa a una de tipo musical, que interpretamos como indicador de buen humor y de buena relación entre los miembros, marcado como **canta** (46).

- (4) DAU[00:43.380 --> 00:44.500] Ayer di un paseo.
 DAU[01:13.400 --> 01:16.500] **ruidos**
 FAT[01:17.400 --> 01:17.500] **ruidos**
 DAU[01:18.400 --> 01:27.500] **ruidos**
F1-3

Llama la atención el abundante número de onomatopeyas, pues no son unas categorías léxicas frecuentes en español y su uso suele asociarse en la lengua con el habla infantil. Sin embargo, tienen una notable representación en nuestro corpus.

- (5) MOT[02:15.480 --> 02:18.620] Salen unos filos cuadrados. Porque las pongo en la bandeja y eso lo pongo en la harina. Y en la harina hago así y ahora *crack crack crack crack*. *F2-5*
- (6) FAT[00:11.020 --> 00:55.200] A mí cuando xxx de eso xxx, que de vez en cuando hacen *pipipín pipipín*. *F3-5*

- (7) FAT[14:58.960 --> 15:10.340] *Buam, bu, buam, bum, fum, fum, fum, pan rallado, chucuchú chucuchú, chucuchú, vuelta, chuchuchú, chuchuchú, chucuchú, vuelta. Mamá chucuchucuchucuchum, harina, trucutrucutrum.*
Y como mamá iba más rápida que yo... *F2-5*

4.2. Nivel morfológico

En la morfología flexiva no se observan ejemplos desligados de la norma coloquial, ya que las discordancias relativas al género, al número, a la flexión temporal, modal o aspectual y a la flexión de caso son habituales y propias de la inmediatez y la falta de planificación discursiva. Lo que sí resulta llamativo es la descompensación evidente entre familias, llegando a haber una diferencia del 21,9% entre la familia con menos irregularidad flexiva (9,9% del total) y la familia con una irregularidad flexiva mayor (31,8% del total). Este resultado podría evidenciar una diferencia de tipo diastrático. Si se quisiera argumentar que es el contexto familiar el causante de tanta vacilación e irregularidad, tendríamos que contrastar los datos de las familias y de los miembros con otras muestras de conversaciones informales con conocidos y desconocidos, para advertir si cometen la misma cantidad de irregularidades flexivas. De esta forma, se podría argumentar que la vacilación de la morfología flexiva es un rasgo más frecuente en el familecto y sus causas podrían deberse a la relajación del lenguaje en el contexto familiar y al interés por la transmisión del contenido a expensas de la composicionalidad.

En la morfología derivativa, los mecanismos de derivación y de composición resultan rentables para generar nuevas formas léxicas y neologismos en español. En el corpus destaca la derivación superlativa en los padres asociada al lenguaje o argot juvenil (cf. Montero Curiel, 2011), como resultado de la confluencia de miembros de diferente nivel sociocultural (§1); así como la abundante presencia de sufijos apreciativos, que se adicionan a cualquier categoría léxica y en cualquier contexto y condición, p.e. sobre el primer elemento del compuesto (16) o en concordancia con el sustantivo sufijado (13-15).

- (8) FAT[07:37.040 --> 07:39.040] Y lo que pasó el miércoles no lo entiendo, porque en el ajarafe había un *atascón* también de narices. *F2-5*
- (9) MOT[01:58.040 --> 01:59.520] A mí me pasa eso con las migas. Las migas me encantan, pero es una tapita. Como me coma un *platazo* ya me satura. *F2-4*

- (10) MOT[03:22.480 --> 03:26.940] El empanado queda *súper crunchy*. *F2-5*
- (11) SON[01:26.000 --> 01:28.340] Porque eso en verdad es *sanote*, *¡eh?* *F2-4*
- (12) DAU[03:54.000 --> 03:57.980] Allí hay un *pueblucho*. *F1-4*
- (13) FAT[04:32.200 --> 04:38.200] Se estará tomando su cafelito antes de entrar en su *pisito*... *turistiquito*. *F1-1*
- (14) DAU[21:23.900 --> 21:27.260] Tengo ganas de coger un *poquito* de *solito*. *F1-5*
- (15) FAT[14:04.180 --> 14:05.700] Un baño, eso, en la piscina no. Y luego la parte que llega hasta el naranjo de la casetilla, hacer ahí una especie de *cuartito grande*. *F3-5*
- (16) FAT[14:12.160 --> 14:13.160] Ya empieza el *jamoncito-York*, luego a la plancha... *F4-1*

Otra categoría interesante en relación con la sufijación, y contemplada en la caracterización del familecto, es la presencia de sufijos con bases inesperadas, que genera derivados sobre formas ya coloquiales (22-23), juego léxico (17) y formas morfológicas no previsibles según las reglas morfológicas del español (18-21).

En particular, destaca la aparición de una suerte de concordancia entre sufijos apreciativos (*pisito turistiquito*, *un poquito de solito*, *cuartito grande*), que podría englobarse en la categoría de lo que Bajo Pérez (2020) denomina *morfolología excéntrica*, y que posiblemente constituye, a diferencia del uso regular de la derivación apreciativa, y al igual que formaciones como *soledosa* (17) o *coquetamen* (20), una particularidad del familecto respecto a la variedad coloquial, por su marcado carácter lúdico y presumiblemente irónico.

- (17) DAU[02:35.680 --> 02:37.920] En la soledad más *soledosa*. *F1-2*
- (18) SON[03:48.500 --> 03:49.500] ¿Tan *levatorizado*? *F2-1*
- (19) FAT[13:54.960 --> 13:55.960] ¿El *kebaquero*? ¿Y ha cerrado el bar? *F2-2*
- (20) MOT[14:26.260 --> 14:27.820] *Croquetamen*. *F2-5*
- (21) SON[00:56.000 --> 01:02.000] Que dicen que en España, por lo visto, estaba prohibida y no se pudo pintar *cocheres*. *F5-2*
- (22) DAU[03:59.540 --> 04:02.540] Juan es igual, mira. Tiene un *napión*. *F1-2*

- (23) SON[02:09.960 --> 02:14.720] Cuando os ponéis un *cacharraco* de salmorejo ahí con 20 cosas, a mí me parece también un poco desagradable. *F2-4*

La parte más interesante del análisis del plano morfológico radica en la notable cantidad de compuestos neológicos, que se puede defender como producto del juego léxico entre los integrantes de las familias. La composición es un procedimiento de formación de palabras considerado poco productivo en español (Val Álvaro, 1999, p. 4759), por lo que una presencia tan significativa podría ser un indicador sociolectal. Cabría realizar un análisis más pormenorizado de estos neologismos, ya que los límites entre las palabras compuestas y otras estructuras gramaticales como las locuciones, las voces prefijadas y las construcciones en aposición son polémicos (Marqueta, 2017; 2022).

- (24) FAT[20:33.200 --> 20:34.200] Ya se han ido los ojos.
 DAU[20:34.200 --> 20:34.200] ¿Qué?
 FAT[20:35.200 --> 20:37.200] *Los hombres ojo. F1-1*
- (25) DAU[05:18.360 --> 05:22.560] Las pizzas están muy buenas
 y no están tan caras, tipo es *precio pizza. F1-3*
- (26) FAT[06:21.600 --> 06:24.660] ¿No había palitos *remuevecrepes?* *F1-3*
- (27) MOT[11:19.120 --> 11:21.120] Da igual que me lo digas.
 ¿El *viejo joven* quién es? *F2-3*
- (28) FAT[03:42.940 --> 03:44.900] Bueno, si tú ya tienes que *medio-empezar* el viaje.
 Tú te vas el lunes. *F2-4*
- (29) DAU[07:01.040 --> 07:04.540] ¿El *loro grande?* Eh, el *papagayo-guacamayo.*
 Es que, no, hemos buscado papagayo. *F3-2*
- (30) SIS[13:22.080 --> 13:44.380] Y dice que a lo mejor en toda la noche habían
 puesto nada más que *música Erasmus. F4-2*
- (31) SON[05:11.980 --> 05:12.980] El *cura narco. F5-2*

4.3. Nivel sintáctico

En el plano sintáctico se comprueba que la creencia de que la sintaxis del español coloquial es sencilla y poco refinada es infundada (§2.1), pues se observan numerosas subordinadas y diversidad de construcciones. El plano sintáctico es posiblemente

el que más similitudes guarda con el estilo coloquial, como así notifican los casos de anacolutos (32), de enunciados suspendidos (33), de variedad de órdenes de constituyentes (34-35), la variedad de funciones de los nexos coordinantes y subordinantes (36), la concatenación sintáctica (37), la sintaxis parcelada (38) y los usos de la deixis personal y la impersonalización de la primera y la segunda personas (39).

- (32) FAT[07:26.880 --> 07:28.960] *O sea, Haku no creo que sea capaz de cogerla.*
Extrañamente. No, porque cuando ya era chica se fue tirado para ella.
A no ser que fuera una lagartija. Con la cabeza y cuando le xxx, eh, eh, toqué. F3-5
- (33) FAT[11:53.040 --> 11:55.040] *Que por la mañana los semáforos los vas pillando todos en verde. Y había tal cantidad de coches metidos en la avenida La Paz...*
Luego, ya una vez llegues al hotel NH, ahí ya no hubo atasco. Y, curiosamente, en todo lo que era avenida San Pablo tampoco. F2-5
- (34) FAT[04:11.200 --> 04:16.660] *Yo el jueves no hay problema porque yo me voy por la mañana, pero vengo aquí a comer y después llevo a mamá a la consulta. F2-4*
- (35) FAT[00:20.720 --> 00:35.620] *Mamá, gel de baño, hay que comprar. Nescafé, bote grande, más grande que este, hay que comprar 3x2. Titi, yogures, hay que comprar 3x2, Activia. F2-2*
- (36) FAT[06:55.960 --> 06:57.900] *y después coches de flota yo no quiero y me gusta saber de dónde viene el coche, ¿sabes?. F4-4*
- (37) FAT[17:50.380 --> 17:57.640] *Como yo espere el domingo, la comida, para cuando yo llame el sábado para reservar para el domingo, no queda un puto sitio para comer. Si, además, encima es pequeño. F2-5*
- (38) MOT[14:20.720 --> 14:24.480] *El pan lo están haciendo ahora. Por la tarde no trabajan y ahora hacen la masa de pan y a las 8 de la mañana está el pan hecho, pero los panaderos trabajan de noche. F2-1*
- (39) FAT[03:14.760 --> 03:40.640] *Hombre ha, bueno, antes de venir para acá a las cinco o a las seis que se llamó. F3-1*

De los rasgos esotéricos previstos se observan ejemplos de interrogativas *in situ* (40) y de construcciones antipasivas (41). Como rasgos novedosos, se hallan casos de cambio de la diátesis verbal (42) y de elisión de categorías funcionales (43), aunque haría falta un estudio exhaustivo comparado para descartar la presencia de estos rasgos en el lenguaje coloquial.

- (40) MOT[19:00.580 --> 19:12.380] *¿Y a Dos Hermanas por dónde, María? ¿Por la universidad xxx para delante?*. *F1-5*
- (41) FAT[19:51.460 --> 19:55.500] *Un miércoles yo podría comerme mi menú ejecutivo si no vengo a casa.* *F2-5*
- (42) FAT[04:00.440 --> 04:20.720] *Había una señora del otro, de la calle, esta abajo, de Castilleja, quejándose cómo está aparcado el coche.*¹¹ *F3-5*
- (43) MOT[15:05.120 --> 15:07.120] *Un pedazo o castillo.* *F1-4*
- (44) FAT[17:40.600 --> 17:46.360] *Madre compró el sábado una berenjena para hacerla frita. Ahí está muerta o risa.* *F1-2*

4.4. Nivel léxico

El plano léxico es, junto con el morfológico, el más rentable para observar diferencias con respecto al español coloquial. De las características planteadas por Gaviño (2023) solo se registran el uso de proformas o palabras comodín (*guay, hacer, coger, hombre, eso, etc.*) y de redundancias léxicas (45, 46).

- (45) MOT[12:30.660 --> 12:35.320] *Yo digo, hoy no hablo con ella, vamos. Llevo toda la... La semana pasada no la vi en toda la semana.* Y esta semana... *F4-3*
- (46) DAU[02:26.200 --> 02:32.200] *canta* *Esos dos llevan los mismos botines en los pies.* *F1-1*

Como era de esperar, la mayoría de vocativos empleados designan el rol familiar que sustenta cada miembro de la familia (*mamá, papá, madre, padre, hija, niño*). Entre los hipocorísticos registrados se observan nombres propios con sufijación apreciativa (*Carlitos, Anita*), habituales para mostrar afectividad; apellidos determinados (*el Ramírez*) y mote de origen idiosincrásico (*Manona, boli, lil, Juan*)¹².

11. Como indica Bogard, este es un comportamiento sintáctico en el que el verbo ya no es capaz de proyectar la función de objeto, al igual que sucede con las construcciones pasivas y anticausativas, de ahí el estatus de oración intransitiva (1999, p. 306).

12. Todos los apodos cuentan con una amplia trayectoria de uso y nacen de anécdotas familiares concretas que sus miembros reconocen. *Manona* es un compuesto procedente del sintagma *mamá Lola*, resultado de la síncopa de la vocal final de *mamá* y de la lateralización de la bilabial nasal tras desaparecer la alveolar líquida; todo ello como consecuencia de la dificultad de pronunciación del habla infantil. *Boli* procede de acortar el sustantivo derivado *bolilla*, con el que la familia apodó a uno de sus miembros debido a sus características físicas en un momento concreto. *Lil* es una abreviatura del inglés *little* (pequeño) con la que se hace referencia a la estatura o edad de una persona; en el caso de esta

- (47) DAU[04:50.700 --> 04:52.940] Nada, *padre*, te haces el cojo o algo. *F4-1*
- (48) FAT[17:40.600 --> 17:46.360] *Madre* compró el sábado una berenjena para hacerla frita. Ahí está muerta risa. *F1-2*
- (49) FAT[01:24.400 --> 01:30.500] ¿Y hoy no tendrás nada para esta tarde, no, *hija*? *F1-3* FAT[01:24.400 --> 01:30.500] ¿Y hoy no tendrás nada para esta tarde, no, *hija*?
- (50) FAT[01:35.900 --> 01:41.520] Era a partir del 15, pero es el cumple de *Carlitos* y lo he rechazado hasta el 16. Que el día 15, creo que cae en sábado, ¿puede ser? *F3-3*
- (51) FAT[08:01.580 --> 08:03.580] Tú, cuando te pones así, *Anita*, a mirar te falta un papagayo aquí, ¿eh?, y una pata de palo. *F3-3*
- (52) FAT[04:48.260 --> 04:59.780] En realidad se ha sentado dos veces pero bueno, yo después de leer, legalmente, no tiene por qué devolverme nada. Aunque después de *Manona* debería sacar los dos últimos dos años de cuentas del Corte inglés. *F3-1*
- (53) MOT[06:49.120 --> 06:51.120] Qué fallo ha tenido hoy *el Ramírez*. *F2-3*
- (54) FAT[00:28.680 --> 00:29.120] Ea, *boli*. Para ti y para Pablo. *F2-5*
- (55) MOT[16:08.420 --> 16:10.980] ¿*Lil*, que vas a querer para tu cumple, eso? Que se acercan días. *F4-1*
- (56) DAU[08:27.200 --> 08:29.200] ¡*Juan*, viento!. *F1-1*
- (57) FAT[09:33.860 --> 09:34.520] Qué asco, *Juan*. Queso caliente. *F1-3*

El corpus abunda en dialectalismos (58-59) y extranjerismos (60-61), que reflejan las marcas sociolingüísticas de los hablantes. La comunicación se empapa de formas infantiles (62-65), de léxico afectivo (66-69), pero también de expresiones juveniles, especialmente entre los padres (70-74) y de tecnicismos (75-79). Esto es una prueba fiel de que en la conversación familiar convergen muchos tipos de argots distintos, resultado natural de la variedad de perfiles que conviven en el hogar familiar. Dentro del léxico soez, se observan palabras malsonantes, empleadas a modo de queja personal (80) o dirigidas con intención hiriente a personas externas al núcleo familiar, pero nunca a otros miembros (81).

familia, para describir a uno de sus miembros. *Juan* es un nombre propio que reconocemos como mote, porque es usado indistintamente por la hija y el padre de la Familia 1. Su origen se encuentra en uno de los personajes del carnaval de Cádiz, que la familia asumió para nombrarse entre sí.

- (58) MOT[03:00.500 --> 03:10.500] Va súper bien el aceite del árbol del té para la tiña, ¿eh? Pero vamos, se te ha quedado el cuello *despeluchado* entero, ¿eh? *F2-1*
- (59) FAT[01:38.440 --> 01:53.760] Pues prepárale un cubo, un poquito, un palaustre, con un poquito de agua, lo pones líquido, pero no *cardivache*, *cardivache*, sino líquido. *F5-3*
- (60) MOT[03:22.480 --> 03:26.940] El empanado queda *súper crunchy*. *F2-5*
- (61) DAU[03:40.180 --> 03:41.840] Los *homeless*, claro, eso. *F3-4*
- (62) FAT[13:38.940 --> 13:39.940] Voy a hacer un *pis*. *F3-4*
- (63) DAU[12:50.780 --> 12:52.660] Uy, me he tirado un *pedo* horrible, Lil. Voy a bajar aquí porque es horrible. *F4-3*
- (64) DAU[10:52.420 --> 10:52.940] Está *repe*, ¿eh?, ¿me cambias el *cromo*? MOT[10:53.980 --> 10:54.340] Está *repe*. Quiere otro *cromo*.» *F4-4*
- (65) MOT[06:12.540 --> 06:21.540] Nadie quiere quitarle el *chupe*. Yo, como no lo quise... Y ustedes tampoco. Pues nada. Pero... *F5-4*
- (66) MOT[04:25.000 --> 04:30.980] ¿"Feos"? A ver, ¿qué quiere *mi niña*?, ¿Un peluche? *F1-4*
- (67) MOT[05:16.120 --> 05:18.120] No, del pan y de otras cosas, *cariño*. *F2-3*
- (68) MOT[13:28.180 --> 13:37.780] *Te quiero*, Antonio. *F2-5*
- (69) DAU[05:13.160 --> 05:20.780] Espérate, Haku, vente para acá, *cariño* mío. Que ya ha puesto papá esto, porque vamos... De verdad, cómo es, ¿eh? *F3-3*
- (70) MOT[06:11.180 --> 06:22.240] ¡"Organicé a María!" Es una *máquina*. *F2-4*
- (71) FAT[06:39.040 --> 06:41.040] Yo lo que te digo es que el otro día para salir hacia la Avenida de la Paz, la zona esta de aquí de Montequinto, el semáforo y *todo el rollo* estaba de coches...*a lo bestia*. *F2-5*
- (72) FAT[06:55.040 --> 06:57.040] Ya iba con *un mogollón de coches*. *F2-5*
- (73) FAT[08:19.040 --> 08:37.040] Se pone *a reventar*. *F2-5*
- (74) FAT[03:21.280 --> 03:23.100]...el artilugio. ¿A qué hora llegó *el nota*? *F3-3*

- (75) FAT[12:18.180 --> 12:24.020] Estamos utilizando tres nada más, porque el que tengo de *almendrodo*, hay quitar esa parte seca. Y tenemos... Y *riego porcote* también. Son seis y tenemos tres. Entonces, utilizaba cinco como máximo y... Entonces, si yo pongo uno de *electroválvulas* como un circuito, donde tengo unos dispersores de estos típicos de invernadero que hace...riegan... *F3-5*
- (76) FAT[00:00.000 --> 00:22.720] Si está bien asentado, el *rasillón* no lo puede poner apoyado en dos puntos ahí, sino que eso tiene que estar una capa de mezcla, el *rasillón* encima asentado en toda la mezcla. Y luego encima del rasillón tiene que haber otra capa de pegamento de *mezcla de mortero* para poner la placa encima. *F5-3*
- (77) SON[06:30.040 --> 06:32.040] Que hizo una especie de *tijereta de pico* xxx, ¿te acuerdas?. *F5-1*
- (78) SON[06:50.040 --> 06:52.040] y ahora está tarde, me voy a ir para allá, porque tengo que dejarle preparado los tres *premarcos* de las puertas a este hombre. *F5-3*
- (79) BRO[01:18.360 --> 01:29.860] Que lo del otro día que era toda la torre de cemento, primero, que lleva mucho cemento, que es un coñazo prepararlo, simplemente, no es que valga dinero, pero es un coñazo prepararlo. Y segundo, se ha *cuarteado*. *F5-3*
- (80) FAT[03:55.240 --> 03:58.240] ¿Qué dices? Y antes pusimos papagayo y estaba ahí. *No tiene ni puta idea.* *F3-2*
- (81) FAT[01:34.260 --> 01:54.180] Pero vamos a ver, como te digo, no es solo aquí. Es que lo que te estaba diciendo, lo que más *me toca las narices*... Es que los que más caros cobran es la iglesia. *Eso me toca los huevos*. O sea, vas a Roma y te cuesta dinero... *F4-1*

Son especialmente relevantes los casos de juego léxico que emplean procedimientos como la metáfora (82), la paronomasia (83-85), la tautología (86), la hipérbole (87), y la antítesis (88), entre otros. Los casos de juego léxico conviven con los usos idiosincrásicos (89-93), configurando dos exponentes muy claros de la representación del juego léxico y de la función identitaria.

- (82) DAU[14:13.200 --> 14:15.200] ¡Me estoy mojando por aquí! (Se recuesta en el hombro de su padre)
 FAT[14:17.200 --> 14:19.200] ¿*Yo que soy?* ¿*Un sofá?* *F1-1*

- (83) FAT[11:01.280 --> 11:04.260] *¿La codicia qué es, una inflamación del codo?*
F1-2
- (84) DAU[07:08.060 --> 07:09.320] *El ingeniero usa el ingenio.* *F3-3*
- (85) FAT[11:10.200 --> 11:13.200] *Juan, los chinos vuelven. Están mal orientados.*
DAU[11:16.200 --> 11:21.200] *Pues es gracioso porque vienen del Oriente.*
Deberían estar orientados. *F1-1*
- (86) DAU[05:21.200 --> 05:28.200] *¿Y por qué no han puesto a las tiendas Yellow Banana? Porque las bananas son yellows.* *F1-1*
- (87) DAU[17:55.600 --> 18:00.360] *Más frito no. Ya esto ha sido el frito del día*
FAT[18:01.600 --> 18:02.360] *El súper frito.* *F1-2*
- (88) FAT[12:12.940 --> 12:13.940] *Bueno, en principio... vamos a tomar por aquí.*
En principio y en final, porque no vamos a (quedarse). *F3-4*
- (89) FAT[07:13.040 --> 07:25.040] *Va a darmel *un maje*.* *F2-5*
- (90) FAT[03:47.680 --> 03:54.220] *Si, le meten a todo. Mira ahí se ha dejado una mesa*
*y no se ha parado. Nos habrá visto cara de *monguel*.* *F1-3*
- (91) DAU[12:00.620 --> 12:02.860] *Sí. Donde tenemos la... *La selva. La selva de Haku.**
FAT[12:04.860 --> 12:05.640] **La selva de Haku.* F3-5*
- (92) SON[17:42.960 --> 17:44.960] *Un *piripiropi*, un *piripiropi*.* *F2-2*
- (93) SON[00:00.000 --> 00:03.100] **La pinga me tula**
MOT[00:03.000 --> 00:05.100] **La pinga me tula.* F2-3*

5. DISCUSIÓN

Tras revisar los datos lingüísticos del corpus elaborado, concluimos que las evidencias acumuladas permiten avalar la hipótesis de que el familecto debe entenderse como una variedad sociolingüística diferenciada. A ello contribuye el hecho de que la familia es un contexto suficientemente distintivo. Las interacciones en el seno familiar son muy diferentes a las que se observan en el resto de interacciones informales, ya que entre los miembros de la familia no hay una relación social o funcional de igualdad que sí se da entre los interlocutores de una interacción informal, sino una jerarquía, que provoca, por ejemplo, que se prescinda de insultos, palabras malsonantes con uso despectivo, faltas de respeto entre los miembros o de

un léxico grosero o soez. La unidad social se preserva, además, mediante códigos diferentes a los de otros grupos informales (funciones identitaria y lúdica) y la complicidad se canaliza a través de otras estructuras y funciones lingüísticas (p.e. aspectos psicosociales de importancia en otras interacciones informales, como la imagen o el posicionamiento, quedan al margen en el caso de los familectos)¹³. Al mismo tiempo, hemos observado que no siempre los rasgos encontrados coinciden con los previstos por Benítez-Burraco y Felíu-Arquiola (2024). Esta circunstancia no significa tanto que el familecto no sea una forma de interacción esotérica, sino, más bien, que no es tan sencillo hacer una traslación de fenómenos de interés para el estudio de la variación tipológica de las lenguas al estudio de la variación de una lengua en particular.

Observamos que el nivel fónico es el que menos coincide con los rasgos esperables del familecto según la propuesta de Benítez-Burraco y Felíu-Arquiola. No se observan casos intencionados de modificaciones secundarias de sonidos habituales, ni de alteraciones prosódicas o de la fonotaxis habitual. Algunos de los ejemplos de juego léxico sí que emplean estrategias fónicas como rimas (*soledad soledosa*) o paronomasias (*oriente, orientados*). Sin embargo, no cabe descartar que algunos de los fenómenos no atestiguados aparezcan en futuras grabaciones familiares, ya que, si atendemos a nuestra intuición como hablantes, es muy posible que los hijos sobre todo acudan a palatalizaciones (*chí* por *si*) o a formas silábicas complejas (*síp* por *si*) como una estrategia de infantilizar su discurso (§2.2), a veces con fines ilocutivos concretos. En contrapartida, sí hemos observados usos fonológicos y entonativos distintivos que no se habían contemplado en la propuesta inicial.

Junto con el morfológico, es en el plano léxico en el que se observan más diferencias con respecto al español coloquial. Además, aquí sí se encuentran muchas de las características previstas para los familectos en la propuesta de Benítez-Burraco y Felíu-Arquiola, pero también observados por otros autores que se han ocupado del estudio de las variedades familiares: compuestos neológicos, sufijación apreciativa excéntrica, superlación típica del lenguaje juvenil, vocativos afectuosos, apelativos particulares, sufijos aplicados a bases inesperadas, o modismos y palabras con significado idiosincrásico. Comprobamos que los mecanismos de creación léxica utilizados en el caso de los núcleos familiares son distintos a los habituales en el español coloquial (composición frente a derivación, abundancia de juegos fónicos y léxicos de diferente tipo). Al mismo tiempo, los significados

13. Cf. Calsamiglia y Tusón, 2001 para más información.

idiomásicos casan con la escasa transparencia semántica y el elevado grado de idiosincrásicos habituales en las lenguas esotéricas, además de sugerir una fuerte intimidad y complicidad, en la línea de lo apuntado por Clancy (2016). En relación con la imitación del habla infantil (*vid. Gordon, 2009*), se observan formas propias de este tipo de lenguaje, usadas tanto por los hijos como por los padres, aunque no hay indicios de procesos fónicos y/o fonológicos que infantilicen el discurso, salvo los ruidos emitidos por el padre y la hija de la primera familia. En cuanto a las características de intertextualidad y de ventrilocuización, corresponden más bien al plano discursivo, si bien los juegos de palabras que van completando los intervinientes durante el diálogo apuntan a un resultado positivo en este sentido.

El caso del nivel sintáctico es quizá el que necesita de una mayor revisión, ya que los rasgos encontrados, que se corresponden en general con los previstos para los familectos por Benítez-Burraco y Felíu-Arquiola, coinciden también en general con la sintaxis de la oralidad y, en concreto, del español coloquial (Vigara Tauste, 1992). No obstante, cabría esperar que sí existiesen diferencias cuantitativas, por lo que en el futuro se pretende comparar las frecuencias de uso de los fenómenos sintácticos de interés entre los corpora familectales y coloquiales.

De esta forma, los resultados obtenidos tras el análisis se resumirían tal y como se contempla en la Tabla 2.

NIVEL FÓNICO	
a) Predominio de onomatopeyas	b) Rasgos fonológicos y entonativos propios
NIVEL MORFOLÓGICO	
a) <i>Mayor vacilación de morfología flexiva</i>	b) Compuestos y derivados neológicos
c) Sufijación de base inesperada	d) Abundante sufijación apreciativa
e) Superlación juvenil entre los padres	f) Presencia de morfología excéntrica
NIVEL SINTÁCTICO	
a) Presencia de interrogativas <i>in situ</i>	b) Presencia de construcciones antipasivas
c) <i>Cambio de diátesis verbal</i>	d) <i>Elisión de categorías funcionales</i>

NIVEL LÉXICO	
a) Vocativos familiares, hipocorísticos y apodos, algunos de carácter idiosincrásico	b) Presencia de formas infantiles y de léxico afectivo
c) Empleo de tecnicismos	d) Uso de expresiones juveniles y de extranjerismos entre los padres
e) Procedimientos de juego léxico	f) Palabras de significado idiosincrásico
g) Empleo de palabras malsonantes sin intención de regular los acercamientos sociales	

Tabla 2. Propuesta de caracterización de los familectos a partir del análisis de los datos del corpus. En cursiva se presentan aquellos fenómenos que aún no podemos asegurar

Los resultados muestran además que, en el ámbito familiar, la estrecha relación vivencial y el marco discursivo íntimo producen rasgos específicos, especialmente en los niveles morfológico y léxico, debido principalmente a la relevancia que cobran las funciones identitaria y lúdica. Quisiéramos, por tanto, que nuestro estudio anime también a replantear la caracterización (y la categorización) tradicional del estilo coloquial, cuando menos, para reconocer que la diversidad funcional en la comunicación puede traducirse en la generación de variedades lingüísticas sustancialmente autónomas, y que este proceso se ve facilitado por contextos de alta intimidad. En este sentido, si bien los modelos de variación lingüística han contemplado habitualmente la coloquialidad como un factor relevante, creemos que no han abarcado adecuadamente todas las posibilidades de la lengua hablada. Por ejemplo, en el modelo variacional de Koch y Oesterreicher (1990/2007) al que hemos referenciado anteriormente, el familecto se situaría en el polo de la inmediatez comunicativa (§2.1). Sin embargo, aunque este modelo ofrece una estructura útil para situar este y otros tipos discursivos en un espacio de variables contextuales, el concepto de familecto aporta matices adicionales que escapan a esta propuesta. Una razón es que la comunicación familiar no solo se define por las condiciones situacionales inmediatas, sino también por una funcionalidad comunicativa específica y una identidad sociolingüística que se desarrolla a lo largo del tiempo dentro del grupo. Son estas características las que confieren al familecto rasgos idiosincrásicos (en particular, morfológicos y léxicos), que podrían no estar completamente interpretados por el enfoque clásico del *continuum* y sugieren, por

ello, la necesidad de incorporar dimensiones adicionales, entre otras, las relacionadas con la identidad y la interacción prolongada.

6. CONCLUSIÓN

En este artículo se han sentado las bases, sobre todo metodológicas, para el estudio de los familectos en español. El sistema de trabajo desarrollado, las herramientas implementadas para ello y los datos encontrados invitan a seguir investigando estas formas de comunicación singulares y que tanto pueden aportar al conocimiento de la variación intralingüística del español.

Un asunto de especial relevancia que es preciso seguir discutiendo es el estatus del familecto como variedad diferenciada, puesto que muchos investigadores lo siguen considerando una muestra de la variedad oral informal de la lengua. Los resultados de este estudio preliminar sugieren que hay indicios suficientes como para distinguir un sociolecto familiar propio. Ciertamente, el familecto no se puede estudiar sin tener presentes otras variedades de la lengua. En realidad, es un crisol donde se amalgaman lectos conocidos de diferente tipo, determinados por la edad, el sexo, la formación, la actividad y hasta la procedencia geográfica. A esto se suman las dinámicas propias de la conversación. Es impensable, por ello, que el familecto no se solape sustancialmente con fenómenos bien estudiados y establecidos, y que se desvincule del registro oral coloquial e informal. Sin embargo, también confluyen en él otros elementos más restrictivos y menos estudiados, como el vocabulario íntimo, el juego infantil (y en general, el discurso familiar cooperativo), o las estrategias de negociación de las jerarquías. Hay toda una panoplia de elementos de la comunicación familiar que se refieren al espacio familiar, que no es del todo coincidente con el espacio (público) informal (como los sustantivos de rol familiar, los hypocorísticos, los nombres de mascotas o los sustantivos que refieren espacios de la casa y objetos habituales). Su semántica es también atípica, con una motivación y significación que se destaca por su carácter imprevisible y altamente expresiva, de una naturaleza muy peculiar, y que tiene como finalidad principal crear esa visión del mundo única que es propia de cada familia.

Entre las líneas futuras de la investigación a este respecto, quedaría por refinar la caracterización que proponemos para el familecto. Para ello, pretendemos ampliar el número de familias analizadas, pero también cuantificar los fenómenos observados. En concreto, buscamos comparar la frecuencia de los rasgos registrados en nuestro corpus que puedan ser compartidos con otras variedades con la frecuencia que

puedan tener en tales variedades. Por ejemplo, cabría estudiar las características del discurso de los sujetos analizados en nuestro trabajo con su discurso en otros contextos comunicativos formales e informales, como sus interacciones con amigos íntimos o sus manifestaciones cuando conversan en el trabajo. A su vez, creemos que el familecto podría suponer un laboratorio idóneo para estudiar las lenguas esotéricas, a las que se tiene poco acceso en general, y contribuir así a los estudios de tipología sociolingüística, que buscan determinar la relación existente entre la configuración sociocultural de los grupos humanos y las características estructurales de las lenguas que hablan (Trudgill, 2011). En último extremo, si el familecto puede equipararse a las lenguas fuertemente esotéricas, podría constituir un modelo para las lenguas habladas en la prehistoria remota, cuando las sociedades humanas consistían, en esencia, en unidades familiares más o menos extensas.

DECLARACIÓN DE CONTRIBUCIÓN DE AUTORÍA

María Alcedo Ortiz: Análisis formal, Curación de datos, Investigación, Redacción–borrador original, Recursos.

Bárbara Marqueta Gracia: Redacción–revisión y edición, Supervisión, Validación, Visualización.

Antonio Benítez Burraco: Conceptualización, Metodología, Supervisión, Redacción–revisión y edición, Validación, Visualización.

FINANCIACIÓN

Esta publicación es parte del proyecto de I+D+i PID2023-147095NB-Ioo, financiado por MICIU/AEI/10.13039/501100011033 y por FEDER/UE.

REFERENCIAS

- Bajo Pérez, E. (2020). *¿Morfología excéntrica del verbo?: El diminutivo en los pretéritos perfectos compuestos de las coplas flamencas*. Universitas studiorum editrice.
- Beinhauer, W. (1968) *El español coloquial*, Editorial Gredos.
- Benítez Burraco, A. & Felíu-Arquiola, E. (2024) De la variación tipológica a la variación intralingüística (y viceversa): el caso de los familectos. *Revista Española de Lingüística*, 54, pp. 9–28. doi: <https://doi.org/10.31810/rsel.54.1.1>
- Bogard, S. (1999) Construcciones antipasivas en español, *Nueva Revista de Filología Hispánica*, 47, pp. 305–327. doi: <https://doi.org/10.24201/nrfh.v47i2.2103>
- Briz, A. (1996). *El español coloquial: situación y uso*. ArcoLibros.
- Briz, A. (1998). *El español coloquial en la conversación: esbozo de pragmagramática*. Ariel.

- Briz, A. (2010). Lo coloquial y lo formal, el eje de la variedad lingüística. En R. Castañer Martín, & V. Lagüéns Gracia (coords.), *De moneda nunca usada: estudios dedicados a José M.ª Enguita Utrilla*, pp. 125–133. CSIC.
- Briz, A. & Grupo Val.Es.Co. (2002). *Corpus de conversaciones coloquiales. Anejo de la revista Oralia*. Arco Libros.
- Calsamiglia, H. & Tusón, A. (2001). *Las cosas del decir. Manual de análisis del discurso*. Editorial Ariel.
- Canary, D. J., & Stafford, L. (1992) Relational maintenance strategies and equity in marriage, *Communication Monographs*, nº 59, pp. 243–267.
doi: <https://doi.org/10.1080/03637759209376268>
- Chen, S., Gil, D., Gaponov, S., Reifegerste, J., Yuditha, T., Tatarinova, T., Progovac, L., & Benítez-Burraco, A. (2024). Linguistic correlates of societal variation: A quantitative analysis. *PloS one*, 19. doi: <https://doi.org/10.1371/journal.pone.0300838>
- Clancy, B. (2016). *Investigating intimate discourse. Exploring the spoken interaction of families, couples and friends*. Routledge.
- Cortés Rodríguez, L. (1991): *Sobre conectores, expletivos y muletillas en el español hablado*. Ágora (2002) Las unidades del discurso oral. *Boletín de Lingüística*, 17, pp. 7–29.
- Gallego Uribe, S. (2006) *Comunicación familiar: un mundo de construcciones simbólicas y relacionales*, Editorial Universidad de Caldas
- Galvin, K., Braithwaite, D. O., Schrodt, P., Bylund, C. L. (2004) *Family Communication: cohesión and Change*, Allyn & Bacon. doi: <https://doi.org/10.4324/9781315228846>
- Gaviño Rodríguez, V. (2023). *Español coloquial*. Universidad de Cádiz.
- Givón, T. (1979). *On Understanding Grammar*. Academic Press.
- Gordon, C. (2009). *Making meanings, creating family: Intertextuality and framing in family interaction*. Oxford University Press.
- Gottman, J.M., Krokoff, L.J. (1989). Marital interaction and satisfaction: a longitudinal view. *J Consult Clin Psychol*, 57, 47–52. doi: <https://doi.org/10.1037/0022-006X.57.1.47>
- Gregory, M. & Carroll, S. (1986). *Lenguaje y situación: variantes del lenguaje y sus contextos sociales*. FCE.
- Gutiérrez Ordóñez (1997) *La oración y sus funciones*, Arco Libros.
- Herrero, G. (1990). La lengua coloquial: concepto y factores que la caracterizan. *Anuario de Lingüística Hispánica* 6, 255–278.
- Huston, T. L., & Vangelisti, A. L. (1991). Socioemotional behavior and satisfaction in marital relationships: A longitudinal study. *Journal of Personality and Social Psychology*, 61, 721–733. doi: <https://doi.org/10.1037/0022-3514.61.5.721>
- Kay, P., (1977) Language evolution and speech style. En B. G. Blount, B.G. & M. Sanches, M. (Eds.), *Sociocultural Dimensions of Language Change*, pp. 21–33. Academic Press.
- Kempe, V., Brooks, P. J., Gillis, S., & Samson, G. (2007). Diminutives facilitate word segmentation in natural speech: cross-linguistic evidence. *Memory & Cognition*, 35, 762–773. doi: <https://doi.org/10.3758/bf03193313>
- Koch, P. & Oesterreicher, W. (2007). Oralidad y escrituralidad a la luz de la teoría del lenguaje. *Lengua hablada en la Romania: español, francés, italiano* (Trad. de López Serena), pp. 20–42. Gredos. (Trabajo original publicado en 1990).

- Levenson, R. W., & Gottman, J. M. (1985). Physiological and affective predictors of change in relationship satisfaction. *Journal of Personality and Social Psychology*, 49, 85–94. doi: <https://doi.org/10.1037/0022-3514.49.1.85>
- López Serena, A. (2007). El concepto de ‘español coloquial’: vacilación terminológica e indefinición del objeto de estudio. *Oralia*, 10, 161–191.
- Lust, B. C. (2006). *Child Language: Acquisition and Growth*. Cambridge University Press.
- Marqueta Gracia, B. (2017). Prefijos preposicionales y compuestos con preposiciones: dos fenómenos independientes. *Lingüística en la red*, 16, pp. 1–24.
- Marqueta Gracia, B. (2020). La delimitación entre compuestos y aposiciones en español. En M. Martínez Atienza de Dios (ed.), *En torno a la delimitación de determinadas categorías lingüísticas*, pp. 71–90. De Gruyter.
- Montero Curiel, M. L. (2011). Mola mogollón: la superlación morfológica y léxica en el lenguaje juvenil. *Revista de estudios de juventud*, 93, 89–104.
- Narbona, A. (2012). *Los estudios sobre el español coloquial y la lingüística*, XLI Simposio de la Sociedad Española de Lingüística, Universitat de València.
- Narbona, A. (2018). *Sintaxis del español coloquial*. Editorial Universidad de Sevilla.
- Narbona, A., Sánchez, M. P., & Lázaro Carreter, F. (1993). *El español hablado en Andalucía*. Ariel.
- Osovská, I. M. (2019). Lexical-semantic space of the German family discourse. Bialykh, V. D., Didukh, L. I., Yu, M. (eds.), *Efficiency level and the necessity of influence of philological sciences on the development of language and literature: collective monograph*, pp. 123–144. Liha-Press.
doi: <https://doi.org/10.36059/978-966-397-170-4/123-144>
- Roberts, J. (2013). Child language variation. En J. K. Chambers y N. Schilling (eds.), *The Handbook of Language Variation and Change*, pp. 263–276. Wiley-Blackwel.
doi: <https://doi.org/10.1002/9781118335598.ch12>
- Schick, J., Fryns, C., Wegdell, F., Laporte, M., Zuberbühler, K., van Schaik, C. P., et al. (2022). The function and evolution of child-directed communication. *PLoS Biology*, 20, e3001630. doi: <https://doi.org/10.1371/journal.pbio.3001630>
- Seco, M. (1972). *Gramática esencial del español: introducción al estudio de la lengua*. Front Cover.
- Søndegaard, B. (1991). Switching between seven codes within one family—a linguistic resource. *Journal of Multilingual and Multicultural Development*, 12, pp. 85–92.
doi: <https://doi.org/10.1080/01434632.1991.999448>
- Trudgill, P. (2011). *Sociolinguistic Typology. Social Determinants of Linguistic Complexity*. Oxford University Press.
- Val Álvaro, J. F. (1999). La composición. En I. Bosque & V. Demonte (coords.), *Gramática descriptiva de la lengua española*, pp. 4757–4838. Espasa Calpe.
- Vangelisti, A. L. (2004). *Handbook of family communication*. Routledge Communication Series.
- Vigara Tauste, A. M. (1992). *Morfosintaxis del español coloquial*. Gredos.
- Wertheimer, M. (1973). Toward a Phenomenological Psycholinguistics of Multilingualism. En D. Krech (Ed.), *The MacLeod Symposium* (pp. 1–17). Cornell University.

- Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117, 543–578. doi: <https://doi.org/10.1016/j.lingua.2005.05.00>



EL ESTILO NOMINAL EN EL CONTEXTO DE LA LECTURABILIDAD: UNA REVISIÓN EMPÍRICA

Susana AZPIAZU TORRES¹

Universidad de Salamanca

Resumen

El trabajo es una revisión del concepto de *estilo de lengua* y de los efectos de la predicación nominal en español, presentados en trabajos previos, desde una nueva perspectiva, la de la lecturabilidad textual, y con un nuevo enfoque metodológico, basado en el estudio empírico de las reacciones de los hablantes. A partir de una prueba de lectura de textos de diversos géneros textuales elaborada con *Qualtrics*, con predominio en cada caso de una determinada estrategia nominalizadora, se pregunta a los informantes por las dificultades que encuentran en la lectura y comprensión de los textos y se los confronta con los mismos textos facilitados por medio de la inteligencia artificial. Los resultados de la prueba refuerzan la idea de que las métricas de lecturabilidad al uso, basadas en la longitud de palabras y en la densidad léxica, y de complejidad sintáctica, basadas en el número y tipo de subordinadas, no son siempre buenas indicadoras de la dificultad o claridad de los textos, pues no tienen en cuenta elementos morfosintácticos más sutiles vinculados al estilo de cada lengua, como son las diversas categorías gramaticales a través de las cuales se lleva a cabo la predicación.

Palabras clave: estilo de lengua; nominalidad; lecturabilidad; lenguaje claro; inteligencia artificial.

THE NOMINAL STYLE IN THE CONTEXT OF READABILITY: AN EMPIRICAL REVIEW

Abstract

The work reviews the concept of *language style* and the effects of nominal predication in Spanish, as presented in previous studies, from a new perspective, that of textual readability, and with a new methodological approach based on the empirical study of speaker reactions. Based on a reading test comprising texts

1. sazpiazu@usal.es;  <https://orcid.org/0000-0003-2296-5893>.

from diverse textual genres - each characterized predominantly by a distinct nominalization strategy and developed using the Qualtrics platform - participants are invited to reflect on the specific challenges they encounter in processing and comprehending the material. Subsequently, they are presented with the same texts, reformulated through artificial intelligence, in order to assess shifts in textual readability and interpretive engagement. The test results reinforce the idea that conventional readability metrics, based on word length and lexical density, as well as syntactic complexity metrics, focused on the number and type of subordinations, are not always reliable indicators of text difficulty or clarity. These metrics fail to consider more subtle morphosyntactic elements linked to the style of each language, such as the diverse grammatical categories through which predication is carried out.

Keywords: Language style; nominality; readability; clear language; artificial intelligence.

RECIBIDO: 20/03/2025

APROBADO: 17/07/2025

1. INTRODUCCIÓN

En algunos trabajos de hace más de veinte años (Azpiazu, 2000; 2004a; 2004b; 2006; 2013) propuse la existencia de un tipo de estrategias discursivas, de carácter morfosintáctico, que servían para aportar más «nominalidad» a algunos tipos de textos. Se trataba de un conjunto de procedimientos que venían a contrarrestar la predicción a través de verbos para desarrollar una predicción nominal, en la que un sustantivo o una forma nominal asume la tarea de aportar la información proposicional relevante. Las «estrategias de nominalización» eran, según esta concepción, herramientas del sistema afianzadas en la norma de ciertos usos lingüísticos destinadas explícitamente a formalizar una expresión alejada del uso común en español². Para hacer estas afirmaciones nos basábamos sobre todo en nuestra propia intuición y en el contraste con otras lenguas con sistemas de predicción nominal más fuerte y desarrollado, como el alemán, el inglés y, de

2. Tal vez sea preciso aclarar que bajo *nominalización* entendemos, casi siempre, el uso de recursos lingüísticos destinados a introducir más unidades nominales en un texto o un discurso, es decir, a hacer más *nominal* o llenar de *nominalidad* un discurso. *Nominalización* es, por tanto, un concepto activo, mientras que *nominalidad* lo es pasivo o resultativo. Cuando hablamos del *estilo nominal* de una lengua, nos referimos al efecto que produce en dicha lengua, en ciertas circunstancias o siempre, que la predicción se lleve a cabo predominantemente por medios nominales, es decir, a la *nominalidad* resultante (Azpiazu 2004a; 2006). Para una explicación más precisa de la diferencia entre *nominalización*, *nominalidad* y *estilo nominal* remito a Bieñ (2013, cap.2).

manera extrema, el sánscrito clásico. En Azpiazu (2004a) se analizaban cinco de esas estrategias: la composición nominal, la predicción nominal o «nominalización»³ a partir de bases verbales, las perífrasis verbo-nominales o estructuras con verbo soporte, la frase nominal y los adverbios en *-mente*. Aunque de naturaleza algo dispar entre sí, todos estos recursos comparten la función de dotar de protagonismo a las unidades nominales como portadoras de información predicativa, restándoselo a los verbos.

A partir de la observación de estas estrategias y de su función discursiva se llegó a la conclusión de que cada lengua tiene un «estilo de lengua», es decir, una tendencia natural («normal», en el sentido coseriano) a desarrollar unos procedimientos de predicción antes que otros en un tipo de comunicación, digamos, estándar o no demasiado formalizado (Azpiazu, 2013). Para el caso del español, se concluyó que su estilo era más «verbal» que «nominal», de modo que la acumulación de estrategias de nominalización era una manera de «desnormalizar» el discurso. La pregunta, que se discute con cierto detalle en Azpiazu (2013, § 2.), es cuáles son los criterios para decidir cómo es el estilo de una lengua o de otra y, en consecuencia, en qué tipo de textos o discursos cabe encontrarlo. En aquel trabajo se proponían algunas pistas sobre cómo dar con él, pero hoy creo que no se puede ser tan realista: el concepto de *estilo de lengua* es fruto de una intuición subjetiva y personal y, por lo tanto, imposible de demostrar empíricamente. Acceder a él directamente no es factible, pues, como señala Enkvist (1987, p. 136), el estilo de cualquier manifestación lingüística es el resultado de una elección que se ha producido antes de generar el texto:

Cuando oímos o vemos realmente un texto, las elecciones ya se han realizado: la investigación no puede seguir disponiendo de ellas (...). Por eso, la operacionalización del enfoque del estilo como elección plantea problemas muy difíciles al lingüista.

En otras palabras, igual que el estilo de un texto, solo podemos definir el estilo de lengua desde nuestros propios prejuicios sobre lo que entendemos como más o menos «normal» en el uso «común» de la lengua, donde tanto *normal* como *común* son, de nuevo, términos cuyos significados están cargados de decisiones personales.

3. Este uso de *nominalización* no es el que se describe en la nota anterior, sino el que se emplea más comúnmente en la gramática. Es un término mucho más concreto que el anterior y refiere a la acción y el efecto de convertir un verbo en un sustantivo deverbal, generalmente con arrastre parcial de la estructura argumental original (*describir* > *descripción*). Para evitar confusiones, trataremos de limitar al máximo el empleo de este término en este segundo sentido.

Sin embargo, el tema del estilo de lengua y del papel moldeador de ciertas estrategias nominalizantes resurge en el marco de un ámbito de estudio que ha cobrado mucha relevancia en los últimos años: el de la «lecturabilidad»⁴ y el «lenguaje claro». En Azpiazu (2004a) se sostiene en varias ocasiones que el recurso a las estrategias de nominalización produce en español textos más opacos semánticamente, es decir, más ininteligibles, aunque entonces aquella afirmación no pasaba de ser una intuición que solo se justificaba teóricamente. Hoy no es inusual encontrar alusiones a las nominalizaciones como recursos causantes de una baja inteligibilidad de los textos, aunque falta toda la reflexión gramatical asociada a estas afirmaciones, así como la explicación de las consecuencias sintácticas que provoca su evitación en los textos en español.

El concepto de *lenguaje claro* se enmarca dentro de un movimiento transnacional destinado a convertir lenguajes de especialidad especialmente difíciles de comprender para gran parte de la ciudadanía, como el administrativo o el jurídico, en un lenguaje preciso y eficaz, es decir, inteligible (cf. Gómez Belart, Colella y Azzaro, 2023). El movimiento del lenguaje claro ha dado lugar a diversas guías institucionales, algunas de ellas auspiciadas desde la Unión Europea, destinadas a una redacción administrativa más accesible al ciudadano⁵. Su relevancia se describe bien en Camacho Muñoz (2023), donde encontramos también la propuesta de definición que proporciona la Federación Internacional de Lenguaje Claro (IPLF):

4. El término *lecturabilidad* es un calco neológico del inglés *readability*, creado para distinguirlo del equivalente sobre base latina *legibility* («legibilidad») y se refiere, de manera general, al grado de comprensión que ofrece un texto en función de su mayor o menor complejidad sintáctica y léxica, mientras que el término *legibilidad* se emplea para referirse a la cualidad externa, sobre todo visual, que presenta un texto. Así, la *lecturabilidad* tiene que ver con los procesos cognitivos que se le exigen al lector para acceder al contenido lingüístico de un texto (Sigaud-Sellos, 2010), mientras que la *legibilidad* se relaciona más con los elementos ortotipográficos del mismo. Es un término muy vinculado al ámbito de la Didáctica de la Lengua y al estudio del desarrollo de la comprensión lectora de los estudiantes.

5. Para este trabajo hemos consultado cinco de las guías de lenguaje claro disponibles para el español de España, que en lo que sigue citaremos por las siglas de los corchetes: el *Manual del lenguaje administrativo del Ayuntamiento de Madrid* [MAM]; la guía de la Comisión Europea traducida al español, *Cómo escribir con claridad* [CEC]; la *Guía de lenguaje claro de la Generalitat Valenciana* [GGV], el *Manual de estilo de lenguaje claro* del Gobierno de Aragón [MGA] y la *Guía de redacción judicial clara* [GRJ], auspiciada por el Ministerio de Justicia. Las cuatro primeras están orientadas al lenguaje administrativo y la última, más específicamente al lenguaje jurídico, aunque las recomendaciones de todas las guías sobre cómo aplicar el lenguaje claro en la redacción de documentos no difieren sustancialmente. Asimismo, se han consultado las recomendaciones para la redacción clara contenidas en el informe de Gómez Belart, Colella y Azzaro (2023).

Una comunicación está en lenguaje claro si la lengua, la estructura y el diseño son tan claros que el público al que está destinada puede encontrar fácilmente lo que necesita, comprende lo que encuentra y usa esa información.

Como se aprecia, esto no es una definición de qué es el lenguaje claro, sino de sus efectos. Esto es así porque esta definición pretende ser válida para toda lengua y, efectivamente, las consecuencias de entender bien un texto son las mismas para cualquier lector en cualquier lengua; lo que no lo es (no puede serlo) es la manera en que se llega a esos efectos, es decir, cómo debe sustanciarse ese lenguaje claro en cada una de las lenguas del mundo.

Lo que suele entenderse por «lenguaje claro» y lo que entendemos nosotros por «estilo de lengua» no es lo mismo, pero ambos pueden usarse como herramienta de acercamiento mutuo. En este sentido, el concepto de «lenguaje claro» asume parte de los prejuicios y, por tanto, casi todas las dificultades de definición que tenía el «estilo de lengua»: para empezar, ambos son idiomáticos, específicos de cada lengua; ninguno puede definirse a partir de lo que es, sino de lo que no es. De ambos nos preguntamos cuáles son sus características y dónde podemos encontrarlos en estado más o menos puro. Y ambos nos plantean el gran dilema: ¿quién decide, y con qué criterios, qué es más o menos claro o qué pertenece al estilo de la lengua? ¿Quién puede juzgar la eficacia comunicativa de un acto de habla?⁶

No es posible responder a estas preguntas en esta forma, y, sin embargo, como veremos, en el caso del lenguaje claro se han ensayado diversas fórmulas y algoritmos para tratar de hacerlo, e, incluso, se han redactado guías y manuales en las que se presupone que sí hay una respuesta. En realidad, como señala la definición citada antes, al lenguaje claro solo es posible acceder *a posteriori*, para lo cual, si el lingüista pretende salir del círculo vicioso de su propia concepción de «lo claro», debe preguntar sobre ello a los demás usuarios de la lengua.

La perspectiva del receptor no se tuvo en cuenta al abordar los efectos del estilo nominal en español hace veinte años, pero, dadas las similitudes señaladas con el concepto de lenguaje claro, nos planteamos si tiene sentido emplear algunas de las herramientas de medición de lecturabilidad disponibles hoy en día para testar alguna de nuestras hipótesis de entonces. En Azpiazu (2013, p. 44) se sugería que dos maneras de acceder al estilo de una lengua eran el análisis interlingüístico contrastivo o traductológico y la observación histórica de la lengua. Proponemos ahora otra,

6. Cf. Azpiazu (2013, p. 45).

vinculada al lenguaje claro y a la lecturabilidad, como es la observación de la reacción de los lectores ante la lectura de determinadas estructuras morfosintácticas.

Así pues, este trabajo viene a completar algunos de los huecos que quedaron sin explicar de aquellos trabajos sobre nominalidad y estilo de lengua, o tal vez a abrir nuevas perspectivas de análisis que arrojen nuevas interrogantes. En lo que sigue (§ 2.) retomaré la explicación teórica del modo en que las estrategias de nominalización operan en el discurso y me centraré, en concreto, en las que en Azpiazu (2004a) se englobaban dentro del rótulo «predicación nominal»: los sustantivos deverbales y las formas no personales del verbo. Mi hipótesis, ahora como entonces, es que son procedimientos morfosintácticos de opacidad textual, pero en esta ocasión trataré de comprobarlo recurriendo a los lectores. Esto me obligará a revisar críticamente algunos de los criterios que se suelen manejar para medir el grado de lecturabilidad de los textos o al describir qué es el lenguaje claro en español. Trataré de mostrar que muchos parámetros que en las guías del lenguaje claro se toman como axiomas de cómo debe ser dicho lenguaje y que, en general, tienden a identificar «claridad» con «sencillez»⁷, no coinciden necesariamente con el estilo del español en tanto que lengua verbalizante ni favorecen necesariamente la inteligibilidad de los textos. Un ejemplo es el mito de la complejidad sintáctica, que trataremos en el apartado 3. En el cuarto se presentará el método de recogida de datos orientados a comprobar la relación entre estilo nominal y lecturabilidad y en el quinto se procederá a explicar los resultados obtenidos.

2. CARACTERÍSTICAS DEL ESTILO NOMINAL: ABSTRACCIÓN Y CONDENSACIÓN

2.1. La primera cuestión a la que hay que atender es por qué el estilo nominal actúa en contra de la lecturabilidad y, por lo tanto, contra el lenguaje claro. Cuando hablo de «nominalización» en este contexto no me refiero solo a la conversión de verbos en sustantivos abstractos (cf. nota 3), sino, en general, a lo que en parte de la lingüística alemana (Sandmann, 1940; Porzig, 1930-31 y 1951), muchas veces ligada a

7. En Azpiazu (2013, p. 44), al hablar sobre algunas cualidades que debería tener todo estilo de lengua, se dice que «la sencillez ha sido un valor que los gramáticos han reivindicado tradicionalmente para sus respectivas lenguas como sinónimo de claridad y corrección expresivas, pero tampoco ha habido (ni puede haber) un consenso sobre en qué consiste realmente una expresión sencilla, ni siquiera un debate sobre por qué una expresión sencilla tiene que ser necesariamente mejor que otra más compleja». En este trabajo se hará hincapié en que la claridad puede ser un rasgo al que aspire el estilo del español, pero se advierte del peligro de identificar sin más «claridad» con «sencillez».

la tradición de estudios del sánscrito (Jacobi, 1903; Hartmann, 1955), se ha entendido como una tendencia de ciertas lenguas a llenar, en ciertas situaciones comunicativas, un texto de unidades nominales, descartando las verbales. En estas lenguas esta forma de predicar se vincula al desarrollo de la lengua culta, literaria y científica, con una cierta sofisticación. En español es más bien un fenómeno circunscrito a la acumulación de diversos procedimientos morfosintácticos, entre los cuales la predicación por medio de sustantivos deverbales es uno más. Como se ha señalado arriba, en este trabajo nos centraremos, además de en ella, en el uso de las formas no personales del verbo, especialmente, participios y gerundios. Todos ellos confluyen en un mismo punto discursivo y provocan similares efectos, que pueden sintetizarse en dos: la abstracción y la condensación del contenido proposicional⁸.

2.2. Abstracción

La abstracción es la consecuencia lógica de elegir la categoría nominal como forma de predicación preferente frente a la verbal. Como sabemos bien, el verbo es la categoría déictica del español por excelencia: obliga al hablante a situarse él y su discurso en un aquí y ahora, y a organizar su contenido en torno a estas coordenadas. Por el contrario, el nombre impone la objetivización de las referencias o, como señala Iturrioz (1985, p. 401), la «reificación del contenido». Este término lo toma Iturrioz de Porzig (1930-31)⁹, que es sin duda el autor que más énfasis puso en el estudio de la abstracción (o «abstractividad», *Abstraktheit*, en el original, una cualidad que afecta ciertos sustantivos y expresiones lingüísticas) como procedimiento de predicación. Para Porzig, la «abstractividad», lejos de ser una característica dada en ciertos sustantivos, es una manera particular de predicar. Las propiedades semánticas de los abstractos no están tanto en lo que designan como en el hecho de que suponen la transcategorización de un contenido proposicional, es decir, en el mismo hecho de que este contenido se concibe como un objeto. Cuando esto sucede, se abandona toda referencia déictica e individual en el enunciado, o lo que es lo mismo, se generaliza el contenido proposicional. El predicado nominalizado

8. Bień (2013, cap. 3) menciona algunos efectos más de la nominalización del discurso, como el distanciamiento, la estatividad, la ambigüedad, etc. Considero que todos ellos son, en última instancia, consecuencias discursivas de estos dos rasgos principales.

9. «Das echte Abstraktum stellt sich sprachlich-deskriptiv immer dar als Vergegenständlichung eines Satzinhaltes vom Prädikat aus» (Porzig, 1930-31, p. 72 y parecido en Porzig, 1951, p. 146). («El verdadero abstracto siempre se presenta en términos lingüístico-descriptivos como reificación de un contenido oracional que proviene del predicado», traducción propia).

no solo carece de las marcas verbales de tiempo, modo, aspecto, persona, etc. propias del verbo, sino que se desvincula también del momento real de la locución. Cada evento del que se quiere decir algo, cada acontecer, se presenta como un concepto estático, donde se borran los límites entre el proceso y el resultado (Hartmann, 1955).

Iturrioz (1985, p. 400) lo desarrolla del siguiente modo, a propósito de la concepción de frase nominal de Benveniste:

Con el progresivo aumento de nominalidad crece la prevalencia del principio de generalización, se pierde gradualmente la solidaridad con las instancias individuales del discurso; según Benveniste los enunciados con predicado nominal tienen un carácter general, incluso sentencioso, expresa verdades o normas generales y necesarias, no hechos históricos ni situaciones contingentes y actuales. La abstracción sustantiva consiste no solamente en la desconexión del discurso y por tanto en la pérdida de aquellas categorías que garantizan el anclaje del enunciado con las situaciones específicas, sino que requiere además la pérdida de otras propiedades esencialmente verbales como la obligatoriedad de la valencia o la concordancia obligatoria que prestan igualmente especificidad al verbo y con él al enunciado.

Por ejemplo, cuando decimos *No esperábamos la llegada de Ana*, no solo tratamos la acción de llegar Ana como un objeto del que se puede predicar (*No esperábamos eso*), sino que también la desposeemos de marcas gramaticales explícitas, como son el tiempo (no sabemos si Ana había llegado ya o si iba a llegar), el aspecto, la modalidad, etc. Ciertamente, en la expresión abstracta hay un componente relacional y un significado; el problema es que ese significado carece de la textura referencial que proporciona la expresión deíctica del verbo, de modo que se vuelve más interpretable y ambiguo. El receptor, que no cuenta ya con las marcas referenciales que suele proporcionarle el verbo, destinadas a orientarle por el sentido del discurso, debe poner también más de su parte para acceder a la comprensión integral del mensaje.

Por otra parte, dentro de las consecuencias directas de la abstracción encontramos fenómenos de orden estructural que afectan al modo en que se organizan los argumentos. En una predicación verbal los argumentos están bien marcados sintácticamente, sea por medio de casos (acusativo, dativo, ablativo, instrumental), sea por medio de preposiciones o por la posición. Pero en la predicación nominal, los argumentos se adjuntan al nombre de un modo mucho más inespecífico, como simples adyacentes con muchas posibles interpretaciones. Como indica Iturrioz en la cita aducida, la valencia verbal deja de ser obligatoria, de manera que puede o

no mantenerse, pero si lo hace, será de una manera genérica. En la *descripción de Luis* no hay ninguna marca formal que nos ayude a interpretar el papel semántico de *Luis*; la estructura nominal nos deja, de nuevo, «solos» ante una interpretación de la preposición *de* ambigua e inespecífica.

2.3. Condensación

La consecuencia directa de la abstracción es la *condensación* del contenido proposicional, ya que, en definitiva, nominalizar un predicado consiste en sintetizar en un solo concepto la información que cabría en un juicio («mit einem Minimum am sprachlichen Mitteln ein Maximum an Mitzuteilendem», en palabras de Hartmann, 1955, p. 26¹⁰). Esto tiene también importantes consecuencias sintácticas, ya que, al convertirse en una pieza independiente, el predicado nominal se presenta integrado en una oración simple (ver *infra*, ejemplo 1a).

De este modo, el predicado nominal permite aportar nuevas ideas al discurso sin necesidad de recurrir cada vez a una nueva oración, introduciendo en un solo grupo nominal la información de todo un periodo proposicional. Se trata en definitiva de la función encapsuladora de la nominalización, que ha sido señalada, entre otros, por Borreguero (2006 y 2018), González Ruiz (2008), López Samaniego (2015 y 2018) o Nadal (2024). Las relaciones lógicas intraoracionales se establecen ahora directamente o mediante preposiciones (1a y 2a) y no con conjunciones (1b y 2b):

- (1) a. La valiente declaración del acusado *sorprendió a su abogado*.
b. Que el acusado declarara valientemente *sorprendió a su abogado*.
- (2) a. Tras acceder al lugar, *se dirigieron a dos empleadas de la sucursal*.
b. Una vez accedieron al lugar, *se dirigieron a dos empleadas de la sucursal*.

De este fenómeno se ocupó también Lehmann (1988), quien lo denomina «desoracionalización» (*desententialization*), así como Raible (1992), quien, dentro de su teoría de la «junción», o unión sintáctica de unidades oracionales o de estados de cosas (el término lo toma de Tesnière), distingue entre dos polos de conexión: la «agregación», donde las unidades se yuxtaponen sin apenas marcas gramaticales y las relaciones semánticas deben ser interpretadas por el hablante, y la «integración», donde estas relaciones gramaticalizan en forma de casos o preposiciones. En el medio se encontrarían las estructuras en diversos grados de coordinación y de

10. «Con un mínimo de medios lingüísticos, un máximo de información» (traducción propia).

subordinación, donde la junción se lleva a cabo por medio de conectores. Según este modelo, las relaciones predicativas verbales tienden hacia la agregación, pero las nominales se mueven en el polo de la integración.

AGREGACIÓN (<i>yuxtaposición</i> de enunciados)	INTEGRACIÓN (<i>relaciones hipotácticas</i> condensadas en un SN)
El acusado declaró valientemente; <i>el abogado se sorprendió</i> .	Que el acusado declarara valientemente <i>sorprendió</i> <i>del acusado sorprendió a su abogado</i>

Tabla 1: Relaciones predicativas entre los polos de la «agregación» y la «integración»

Al igual que la abstracción semántica, la condensación sintáctica afecta en mayor o menor grado a todas las estrategias de nominalización y, al igual que aquella, exige en sus casos más extremos un esfuerzo interpretativo por parte del receptor del mensaje. Lo que quisiera demostrar aquí es que también es una rémora para la inteligibilidad del discurso, a pesar de que la idea de poder expresar un mensaje con menos palabras o menos subordinadas ha sido considerada, desde los estudios sobre complejidad sintáctica de Hunt (1965; 1970), un criterio de claridad.

3. ESTILO NOMINAL, COMPLEJIDAD SINTÁCTICA Y LENGUAJE CLARO

Para Hunt (1965 y 1970), la complejidad sintáctica se define por el número y variedad de transformaciones que aplica el hablante a la «Unidad Mínima Terminal» o «unidad-T», que viene a ser aproximadamente lo que en la gramática del español consideramos una oración con un único predicado principal (Crespo Allende, Alfaro Faccio y Góngora Costa, 2011, p. 162). Si esta unidad-T, que en principio consta solo de «verbo finito más complementos», desarrolla alguno de sus constituyentes bajo la forma de un nuevo predicado (cláusula subordinada), el resultado es una transformación que conlleva mayor complejidad en la expresión y de madurez sintáctica por parte del hablante. En otras palabras, la madurez sintáctica de un hablante con poca experiencia en la lengua, un niño o un hablante de L₂, se dice, se mide en términos de cantidad de palabras en las cláusulas y en las unidades-T, pero, sobre todo, de cantidad de relaciones de subordinación e interordinación que es capaz de manejar (Hunt, 1970, p. 6; Crespo Allende, Alfaro Faccio y Góngora Costa, 2011, p. 163).

De un modo similar, Salvador Mata (1985) propone hasta cuatro fórmulas para medir la complejidad sintáctica, que se fundamentan en distintos tipos de relaciones

entre las oraciones simples, las coordinadas/yuxtapuestas y las subordinadas. En este trabajo nos interesan dos: una que denomina «complejidad numérica», que consiste en dividir el número de subordinadas por el número de oraciones del texto y multiplicar el resultado por 10, y otra, de «complejidad cualitativa», en la que se tiene en cuenta, por un lado, la naturaleza libre o subordinada de las oraciones del texto y, por otro, el grado o jerarquía de subordinación¹¹.

Según estos dos modelos, las estructuras que se encontraran en los dos polos extremos del binomio agregación–integración de la tabla 1 (*supra*) serían estructuras de baja complejidad sintáctica, pues no hay en ellas hipotaxis o subordinación en sentido estricto. Ahora bien, en la sintaxis moderna se entiende que una estructura nominalizada a través de un sustantivo deverbal es el resultado de una «transformación sintáctica» (con todas las comillas que se le quiera poner a este viejo término de la sintaxis), pues es una construcción que deriva secundariamente de otra, de origen verbal, y hereda más o menos su estructura argumental con una serie de cambios morfosintácticos (Chomsky, 1970; Giorgi y Longobardi, 1991; Escandell, 1995; Picallo, 1999). La cuestión es si, aunque aparentemente la construcción resultante es más breve que la original y carece de marcas de subordinación sintáctica, es también más fácilmente «procesable» o si no es, por el contrario, la manifestación última de un complejo proceso de síntesis y gramaticalización de diversas relaciones hipotácticas¹², que además carece de muchas de las marcas deícticas que aportan las estructuras más agregadas.

En lo que sigue pretendo demostrar que, frente a lo que se afirma en ciertas guías del lenguaje claro¹³, un índice de complejidad sintáctica alto, según las métricas al uso, no implica necesariamente poca claridad textual. De los hablantes nativos adultos de español y con un grado de formación medio puede esperarse que

11. En § 4.2. se aportan más detalles sobre estas dos fórmulas y sobre cómo interpretar sus resultados. Los índices que no utilizamos de Salvador Mata son el de «complejidad estructural oracional» y el de «complejidad estructural proposicional», que vienen a señalar hasta qué punto hay un equilibrio entre coordinación / yuxtaposición y subordinación en los textos. Ambas fórmulas se pueden consultar en Salvador Mata (1985, p. 65).

12. Hunt (1965: 112) entendía también que el uso de este tipo de nominalización (que él denomina «semiproposiciones», *near-clauses*) era también un índice de madurez sintáctica.

13. Por ejemplo, en la GGV, apartados 4.2.5. y 4.2.6., se afirma que las oraciones no deberían tener más de treinta palabras (?) y que hay que evitar encadenar frases subordinadas. La cifra de treinta palabras por oración se repite en Gómez Belart, Colella y Azzaro (2023, p. 321). También en el MGA (p. 34), se habla de «simplificar la sintaxis» y de manejar pocos elementos en las oraciones. Se menciona además una cita de Saint Exupéry: «La perfección no se alcanza cuando ya no queda nada por añadir, sino cuando no hay nada que quitar». En el uso de la lengua, cuando ya no hay nada que quitar, lo que queda es el silencio, es decir, la incomunicación.

sean capaces de manejar sin mayores complicaciones unidades-T complejas, con varios constituyentes en forma de cláusulas subordinadas. Es más, postulo que un hablante de español con una competencia media-normal de su lengua maneja con más facilidad varias de estas estructuras imbricadas, siempre que en ellas el verbo esté conjugado y haya conectores orientadores del sentido, que distintos predicados nominalizados insertos en una unidad-T simple. Hasta donde yo sé, este aspecto no se ha tenido muy en cuenta al tratar los elementos que favorecen o dificultan la lecturabilidad de los textos en español, más allá de algunos comentarios escuetos censurando el uso de los sustantivos deverbales abstractos o de los gerundios de posterioridad, sin más explicación.

4. ESTUDIO EMPÍRICO SOBRE LOS EFECTOS DE LA NOMINALIZACIÓN SOBRE LA LECTURABILIDAD

4.1. *La encuesta Qualtrics*

Como se ha explicado en § 1., al lenguaje claro solo se accede indirectamente, observando los efectos que produce la lectura de distintos textos con diversos elementos susceptibles de generar «complejidad». Es decir, es una aproximación negativa: no podemos dar una definición de qué es el lenguaje claro, pero tal vez sí de qué no lo es. En el caso que nos ocupa aquí, queremos comprobar hasta qué punto los elementos que aportan nominalidad a los textos (sustantivos deverbales y formas no personales del verbo) contribuyen a volver los textos más opacos a la lectura. Para lograrlo, se ha optado por una técnica que empieza a extenderse entre quienes se ocupan de establecer criterios de lenguaje claro: la validación por parte de los usuarios de la lengua, es decir, la consulta directa a los lectores. Para lograrlo, ideamos una breve encuesta con *Qualtrics*¹⁴ en la que se trató de confrontar a una serie de usuarios, por una parte, con textos con abundantes rasgos nominalizantes y, por la otra, con versiones facilitadas de los mismos textos. Dos eran las hipótesis que buscábamos ilustrar con esta prueba:

- a) la condensación sintáctica y la abstracción semántica que conlleva un uso excesivo de unidades nominales complican la comprensión del texto más que los verbos en forma personal;

14. Se trata de una plataforma de gestión que permite recoger y analizar datos, a veces en tiempo real, a través, por ejemplo, de encuestas personalizadas fáciles de generar y difundir. Más información en <https://www.qualtrics.com>.

- b) al menos en una lengua como el español, la lecturabilidad de un texto no debe medirse (solo) en términos de número de palabras (densidad léxica) o cantidad de subordinadas imbricadas entre sí (complejidad sintáctica), pues más elementos léxicos pueden suponer más explicitud semántica y, en consecuencia, mayor claridad e inteligibilidad.

La encuesta se lanzó por correo electrónico y redes sociales (X, Facebook, Instagram y la plataforma académica Studium de la Universidad de Salamanca); reaccionaron a ella en torno a 50 personas, aunque solo una treintena completó la encuesta hasta el final¹⁵. Los informantes tienen casi todos estudios superiores o están cursándolos, son hablantes nativos de español entre los 18 y los 30 años, es decir, individuos de los que cabe esperar una competencia «normal» o «estándar» de la lengua, quizás no siempre especialmente sofisticada pero sí muy asentada y, sobre todo, que están intensamente sometidos a los textos escritos, como autores y, sobre todo, como lectores.

La prueba consistió en leer sucesivamente tres breves textos de tres géneros discursivos tradicionalmente considerados complejos: uno administrativo (A1), uno financiero (B1) y uno judicial (C1) junto con sendas «facilitaciones» encargadas a la inteligencia artificial (A2, B2 y C2, respectivamente). Los textos originales constan de un número de palabras similar (entre 94 y 114¹⁶) y cada uno de ellos se eligió por la abundante presencia en ellos de elementos propios del estilo nominal: sustantivos deverbales en el primero, participios en el segundo y gerundios en el tercero. Para las versiones facilitadas de los textos, se trabajó con ChatGPT¹⁷ y Gemini¹⁸. La razón de trabajar con la IA es doble: por un lado, queríamos mantenernos lo más al margen posible de la manipulación de los textos originales y, por otro, queríamos valorar el grado de acierto de esta herramienta en la consecución de un estilo más

15. El número de respuestas varía con cada texto, porque muchos lectores abandonaron la encuesta hacia la mitad. Los textos A fueron leídos por más de 40 personas, los textos B por 35 y a C llegaron unos 30 informantes.

16. En concreto, el número de palabras todos los textos es: A1 = 94; A2 = 86; B1 = 99; B2 = 88; C1 = 114; C2 = 118.

17. <https://openai.com/>.

18. <https://gemini.google.com/?hl=es>. La elección de Gemini para el segundo grupo de textos se debió al interés por comprobar si los resultados variaban sustancialmente entre ambas herramientas. De hecho, este texto se le dio también a ChatGPT, y se constató que los resultados que se obtenían eran bastante similares.

Así pues, para nuestros propósitos, no pareció muy relevante la elección entre un modelo de IA y otro. Sin duda, una comparación de este tipo puede ser interesante en un trabajo que pretenda analizar el modo en que la IA (o las IAs) lleva(n) a cabo estas simplificaciones y cómo optimizarla(s).

inteligible o claro en español¹⁹. Para medir hasta qué punto el contenido de los textos originales, no facilitados, se comprendía y para mantener la concentración de los lectores durante la lectura, tras estos textos originales había una pregunta de comprensión²⁰.

Después se pedía a los informantes valorar la dificultad de cada texto del 1 al 5, para finalmente, instarles a que valoraran el grado (del 1 al 3) en que ciertos elementos textuales elegidos previamente (nombres abstractos, formas no personales del verbo, formas personales, conectores y puntuación) dificultaban, en su opinión, la comprensión textual. En un caso (textos C), se separó expresamente la valoración del uso de los gerundios del resto de las formas no personales del verbo. La razón de preguntar por los conectores y la puntuación, además de por las categorías verbales y nominales, fue comprobar hasta qué punto las marcas explícitas de subordinación se perciben como facilitadoras de la lecturabilidad y qué sucede cuando los signos de puntuación vienen a sustituir dichas conexiones explícitas. De este modo, presuponemos que una alta valoración de la puntuación como elemento de complejidad textual puede implicar cierta desorientación discursiva²¹, y, por el contrario, que la presencia de más elementos de conexión ayuda al lector a orientarse en el texto, así que los conectores, en principio, serán valorados en términos favorables para la lecturabilidad²².

Por otra parte, una de las grandes ventajas de *Qualtrics* es que permite medir el tiempo de lectura de cada texto. Esta información más «objetiva» (a más dificultades para entender un texto, mayor tiempo de lectura) viene a sumarse a la valoración «subjetiva» que pedimos hacer a cada hablante del grado de dificultad de cada

19. El *prompt* que le daba a la IA era, para todos los textos, «¿Puedes redactar este texto en lenguaje claro?».

20. Dado que sabemos que la comprensión lectora no depende solo de los elementos en los que nos centramos en este estudio y dado que el acierto o fallo en la respuesta a la pregunta de comprensión puede deberse también a cómo se formula la propia pregunta, los resultados obtenidos aquí no han sido considerados muy relevantes para el estudio. Solo se mencionan tangencialmente en algún caso.

21. Hay que señalar que en estos breves textos hay pocas oraciones compuestas o yuxtapuestas. Suelen constar de estructuras complejas en torno a un único predicado principal, de manera que si se eliminan nexos, suelen ser nexos de subordinación. Esto es posible, sobre todo, en contextos de *coalescencia* o correferencialidad sintáctica, que es un recurso integrativo. El término *coalescencia* se toma de Lehmann (1995) y Raible (1992) y se explica en § 5.2.1.1.

22. En Gómez Belart, Colella y Azzaro (2023, p. 104) se habla del peligro de omitir demasiados elementos cohesivos y en algunas de las guías del lenguaje claro (por ejemplo, en GGV, § 4.2.7, y en MGA, pp. 48-51), así como en la herramienta de medición de claridad textual CLARA (ver a continuación), se valora positivamente la presencia de conectores como contribución al lenguaje claro, pues ayudan a visualizar las relaciones argumentativas de manera inequívoca. Si bien esto es cierto, también lo es que las conjunciones van de la mano de verbos en forma personal.

fragmento. La idea era comprobar hasta qué punto ambas medidas estaban en consonancia.

Ciertamente, el diseño de la encuesta no está exento de debilidades que pueden mermar su validez como herramienta de comprobación definitiva. Para empezar, puede aducirse que, aunque pueda parecer lo contrario, la búsqueda de respuestas en otros hablantes no implica un grado de «objetividad» necesariamente mayor que la búsqueda individual de esas mismas respuestas en los trabajos de hace veinte años, pues en definitiva lo que hacemos no es más que elevar la intersubjetividad a la categoría de método empírico (cf. Fernández Pérez, 1984). La medición del tiempo de lectura puede ser, en efecto, una información algo más reveladora, aunque el modo en que se presenta en esta encuesta no esté exento de problemas y, por ello, se empleará solo como dato parcial. Efectivamente, el hecho de que los textos facilitados se presenten cada vez inmediatamente después de los textos originales puede producir un efecto de *priming* que sesgue los resultados sobre la rapidez de lectura (la lectura del segundo texto se acelera porque su contenido acaba de ser procesado en la lectura del primero). Por eso, este factor no lo tendremos en cuenta en este trabajo al medir la diferencia entre textos 1 (originales) y 2 (facilitados), pero sí para medir la diferencia entre los grupos textuales (A, B y C, tanto originales como facilitados), donde el efecto del *priming* ya no opera. En cualquier caso, el objetivo final de la encuesta no es tanto presentar pruebas irrefutables que confirmen los postulados teóricos sobre el efecto de la nominalización en la lecturabilidad de los textos, como aportar datos que puedan enriquecer esas mismas intuiciones y que, acaso, justifiquen en el futuro un estudio de carácter psicolingüístico más robusto metodológicamente²³.

4.2. Contraste con las herramientas de medición de lecturabilidad y complejidad sintáctica

Los resultados de las encuestas de *Qualtrics* se confrontarán también con los resultados que, para estos mismos textos, arrojan algunas herramientas digitales especializadas en medir la claridad o la lecturabilidad, disponibles en la web.

23. Este otro trabajo que se menciona ya ha sido realizado en colaboración con la profesora Olga Ivanova sobre la base de los textos A y C y se encuentra en proceso de redacción. En él se han desarrollado y perfeccionado el trabajo experimental con pruebas de movimiento ocular (*eye tracking*). En esta ocasión, hemos evitado el efecto del *priming* presentando a los informantes los textos originales y facilitados de manera no sucesiva, junto con otros textos de control y aleatorios.

Para este trabajo hemos empleado *Legible*²⁴ (<https://legible.es/>), desarrollada por Alejandro Muñoz Fernández; CLARA²⁵ (<https://clara.comunicacionclara.com/>) de Prodigioso Volcán y *Spanish readability*²⁶ (<https://www.spanishreadability.com/>) (en adelante, SpR).

Legible evalúa los textos según varias fórmulas (Fernández Huerta²⁷, Gutiérrez de Polini²⁸, Crawford²⁹, Szigriszt-Pazos³⁰ / INFLESZ³¹ y lecturabilidad μ de Muñoz y Muñoz³²), que, básicamente, son distintas maneras de establecer relaciones cuantitativas entre el número de palabras de las oraciones o del texto y su longitud en cuanto a número de letras o sílabas. En general, en todas estas fórmulas se considera que, en un rango del 1 al 100, a menor puntuación, mayor dificultad en la lectura. A partir de 50, los textos empiezan a calificarse como «un poco difíciles» o «normales».

Spanish Readability toma las mismas fórmulas de *Legible* y les añade las fórmulas que emplea Bryan Scott para el inglés³³, todas ellas también muy centradas en cuestiones cuantitativas y de densidad léxica. Esta última herramienta hace un balance global de lecturabilidad en términos de edad y grado escolar apropiados para comprender el texto.

Como se aprecia, el tipo de lecturabilidad que miden *Legible* y SpR no es exactamente el tipo que nos interesa: nuestro foco no está en la dificultad en procesar más o menos palabras más o menos largas, sino ciertas estructuras morfosintácticas. Por su parte, CLARA sí resulta más interesante, pues tiene en cuenta aspectos sintácticos más complejos entre los parámetros que evalúa: además de la cantidad de palabras por frase, el uso de palabras comunes y de tecnicismos y la ausencia de referencias a leyes o artículos (es un programa que se emplea mucho para medir la claridad del lenguaje administrativo), evalúa la presencia de conectores, la puntuación, el uso de la voz activa o la pasiva y la sencillez en la construcción

24. Consultas realizadas en enero de 2025.

25. Los resultados corresponden a consultas realizadas el 26/03/2025.

26. Consultas realizadas en enero de 2025.

27. <https://legible.es/blog/lecturabilidad-fernandez-huerta/>.

28. <https://legible.es/blog/comprendibilidad-gutierrez-de-polini/>.

29. <https://legible.es/blog/formula-de-crawford/>. En realidad, en la tabla que devuelve *Legible* no aparecen nunca datos de esta fórmula.

30. <https://legible.es/blog/perspicuidad-szigriszt-pazos/>.

31. <https://legible.es/blog/escala-inflesz/>.

32. <https://legible.es/blog/legibilidad-mu/>.

33. <https://readabilityformulas.com/>.

sintáctica. Al procesar el texto, esta aplicación arroja un porcentaje de claridad, de manera que, a menor porcentaje, mayor opacidad en la lectura, y viceversa.

Además de todo esto, quisimos comprobar cómo de sintácticamente complejos o simples eran nuestros textos según los índices de complejidad sintáctica numérica y cualitativa de Salvador Mata (1985) mencionados en § 3. Como ya se explicó, la complejidad numérica (CN) consiste en dividir el número de subordinadas por el número total de oraciones del texto y multiplicarlo por 10, mientras que la complejidad cualitativa (CC) se calcula otorgando una puntuación a cada tipo de oración (simple, coordinada o yuxtapuesta y subordinada) y a cada rango de subordinación (de primer grado, de segundo, de tercero, etc.) y dividiendo la suma total resultante entre el número de oraciones³⁴. Según Salvador Mata (1985, p. 64), en el caso de la complejidad numérica, el índice resultante mínimo es 10 y no hay un límite máximo, «aunque en la práctica más allá del índice 50 resulta difícil mantener la claridad y la coherencia en la expresión». En cuanto a la fórmula de complejidad cualitativa, el índice mínimo es 1 y tampoco hay límite máximo, aunque en este caso, sostiene, todo lo que supere el 5 implica de nuevo poca claridad y coherencia en la expresión³⁵.

En general, el objetivo de acudir a todas estas fórmulas cuantitativas será evaluar si son métodos apropiados para medir el grado de lecturabilidad de nuestros textos, una vez contrastados sus resultados con las reacciones de los propios usuarios de la lengua.

34. En este trabajo aplicaremos estas dos fórmulas sustituyendo el término «oración» por la unidad-T de Hunt (cf. § 3.), que es una unidad algo mejor delimitada gramaticalmente.

35. En § 3. explicamos que Salvador Mata (1985) propone otras dos fórmulas más, la de «complejidad estructural oracional» y la de «complejidad estructural proposicional», que no hemos aplicado aquí. Son fórmulas que miden la relación entre estructuras coordinadas/yuxtapuestas y subordinadas en los textos. Hemos optado por prescindir de estos dos índices porque, por un lado, nuestros breves textos apenas presentan oraciones coordinadas o yuxtapuestas. Por otro, el motivo de introducir las fórmulas sintácticas solo se debe a nuestro deseo de contrastar algunos índices de complejidad sintáctica tradicionales con los resultados de nuestra encuesta, y eso es algo que creemos poder hacer suficientemente con las dos fórmulas que aducimos.

5. TEXTOS Y ANÁLISIS DE LOS RESULTADOS

5.1. *Predicados nominales deverbales. Textos administrativos.*

5.1.1. Como muestra de un texto con alta concentración de sustantivos deverbales, seleccionamos el siguiente texto administrativo extraído de una antigua normativa de la Universidad de Salamanca (se marcan las construcciones nominales y se numeran para facilitar la comparación con la versión que proporciona la IA):

A1. La interposición del recurso [1] no tendrá efectos suspensivos [2] de la propuesta de nombramiento o formalización del contrato [3], salvo que el Rector, mediante Resolución expresa [4], aprecie que la ejecución del acto [5] pudiera causar prejuicios de imposible o difícil reparación [6] o aprecie en la fundamentación de la impugnación [7] la existencia manifiesta de motivos de nulidad de pleno derecho [8]. A tal fin [9] se formalizará el nombramiento [10] o el contrato con el aspirante propuesto, sin perjuicio de que la eventual estimación del recurso [11] pueda acarrear la extinción automática de la relación contractual o de interinaje del inicialmente propuesto [12].³⁶

A2 es la versión «traducida» al lenguaje claro que ofrece ChatGPT:

A2. El recurso interpuesto [1] no suspenderá [2] el nombramiento o la firma del contrato [3], salvo que el Rector, mediante una resolución específica [4], considere que ejecutar la decisión [5] podría causar daños que no se puedan reparar fácilmente [6] o que vea en el recurso [7] razones claras de nulidad [8]. En caso de que no se suspenda [9], se procederá con el nombramiento [10] o contrato del candidato propuesto, aunque si el recurso se acepta más tarde [11], esto podría hacer que la relación contractual o de interinidad con el candidato inicialmente seleccionado termine automáticamente [12].

Algunas transformaciones son evidentes a primera vista: en al menos cuatro ocasiones ([6], [9], [11] y [12]), cinco si contamos también la transformación de la perífrasis verbo-nominal o de verbo soporte de [2], la estructura nominal original se transforma en un predicado verbal. En un caso [5], el SN se convierte en un infinitivo verbal (cf. *infra* § 5.2.1.1). En otros casos, el sustantivo abstracto se mantiene, pero el SN se simplifica drásticamente [3], [7] y [8].

36. *Resolución de 30 de mayo de 1997, de la Universidad de Salamanca, por la que se resuelve publicar el Acuerdo Normativo por el que se regula el procedimiento de selección de Ayudantes y Profesorado Universitario en régimen de contratación administrativa o interinidad* (BOE-A-1997-12502).

5.1.2. Cuando se les pide a *Legible*, SpR y CLARA que evalúen estos textos, A2 se presenta como claramente más «fácil» o claro que A1. En *Legible* este último se considera «muy difícil» según el test de Fernández Huerta y «árido» según el de Szigriszt-Pazos³⁷ (fig. 1a) y obtiene un 15 % de claridad (es decir, bastante poca) en CLARA (fig. 2a); mientras que A2 es calificado como «difícil» en Fernández Huerta y como «solo» bastante difícil en Szigriszt-Pazos (fig. 1b), y obtiene 43 % de claridad en CLARA (fig. 1b). Por su parte, *Spanish Readability* evalúa en general A1 como «difícil», y lo considera apropiado para un nivel de 2º de Bachillerato (17-18 años), mientras que de A2 dice que es «algo difícil» y baja la edad de comprensión a los 16.

Legibilidad del texto		
índice	valor	dificultad
Fernández Huerta	24.08	muy difícil
Gutiérrez	25.05	normal
Szigriszt-Pazos	19.96	árido
INFLESZ	19.96	muy difícil
legibilidad μ	41.71	difícil

Figura 1a. Índices de lecturabilidad de A1 según *Legible*

Legibilidad del texto		
índice	valor	dificultad
Fernández Huerta	44.4	difícil
Gutiérrez	33.73	normal
Szigriszt-Pazos	39.8	bastante difícil
INFLESZ	39.8	muy difícil
legibilidad μ	45.59	difícil

Figura 1b. Índices de lecturabilidad de A2 según *Legible*



Figura 2a. Índice de claridad de A1 según CLARA

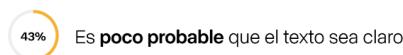


Figura 2b. Índice de claridad de A2 según CLARA

5.1.3. Dado que nuestra búsqueda se centra en los elementos morfosintácticos más que en los puramente léxicos, decidimos analizar algunos factores que podrían influir en la complejidad sintáctica del texto. Los presentamos en la tabla 2:

	Verbos en forma personal	Sustantivos deverbales	Ratio	Nexos ³⁸
A1	6	15	2:5	3
A2	9	9	1:1	8

Tabla 2. Relación entre elementos de predicción verbal y nominal, y número de nexos oracionales en A1 y A2

37. Como se apreciará en todas las figuras de *Legible*, el test Szigriszt-Pazos y la escala INFLESZ emplean exactamente la misma fórmula, pero la valoración de lecturabilidad es diferente en cada una. Una explicación puede encontrarse en <https://legible.es/blog/escala-inflesz/>. A los efectos de este trabajo, preferimos las valoraciones de Szigriszt-Pazos porque matizan algo más el grado de dificultad en los índices de baja lecturabilidad, entre el 0 y el 40.

38. Nos referimos aquí tanto a nexos de subordinación como de coordinación.

Hay en esta tabla dos datos que pueden ayudar a explicar el distinto nivel de complejidad de ambos textos: por un lado, se aprecia el desequilibrio entre predicados nominales y verbales en A1, frente al equilibrio de A2; por el otro, el aumento de predicados verbales va de la mano de un mayor número de elementos conectores. Es lo esperable cuando el estilo del texto se separa del polo de la integración y se desplaza hacia el de la agregación (cf. § 2.3., tabla 1).

Por otra parte, si nos fijamos en el número y la composición de las unidades-T de cada texto, los resultados son los siguientes:

	Unidades-T	Subordinadas	Rango de subordinación	Complejidad numérica - CN (Sub/U-T)	Complejidad cualitativa - CC
A1	2	5	2	25	8
A2	3	8	3	26	10

Tabla 3. Datos de complejidad sintáctica para A1 y A2

Con los criterios sintácticos de Hunt (1965) en la mano, habría que decir que A2 es algo más complejo que A1, pues contiene, en más o menos el mismo número de palabras, una unidad-T más y hasta 8 oraciones subordinadas, frente a las 5 de A1. Además, en A2 hay hasta tres niveles de jerarquización sintáctica, mientras que en A1 no se pasa de un segundo nivel de subordinación. Por su parte, los índices de CN y CC de Salvador Mata dan resultados contradictorios: según el primero, no llegan a ser textos excesivamente complejos, pues no alcanzan el umbral del 50, pero según el segundo, lo son, pues ambos superan el 5. De nuevo, estos índices nos dicen que A2 requiere algo más de madurez sintáctica en el lector que A1.

Pongamos un ejemplo para aclarar esto. Las estructuras [11] y [12] de A1:

- (3) *sin perjuicio de que la eventual estimación del recurso* [11] *pueda acarrear la extinción automática de la relación contractual o de interinaje del inicialmente propuesto* [12],

se traducen en A2 se traduce como:

- (4) *aunque si el recurso se acepta más tarde* [11], *esto podría hacer que la relación contractual o de interinidad con el candidato inicialmente seleccionado termine automáticamente* [12].

El número de palabras de ambas secuencias es prácticamente idéntico (24 vs. 25), pero en (3) todo se organiza en torno a un único predicado verbal (*pueda*

acarrear) cuyos argumentos externos e interno son sendos predicados nominales (*la eventual estimación del recurso* en [11] – *la extinción automática...* en [12]), al estilo de lo ejemplificado en (1a). Por su parte, en (4), [11] es una condicional, cuyo contenido se retoma anafóricamente en *esto*, sujeto de la proposición principal, y [12] es una subordinada sustantiva. Las funciones sintácticas de las dos estructuras se replican (sujeto – CD), pero la construcción de (4) implica introducir nuevos conectores y diferentes jerarquías de subordinación (5):

- (5) *aunque*³⁹  *[si se acepta el recurso]*_{prótasis}, *[esto podría hacer]*
*[que la relación termine]*_{sub. sustant.}_{apódosis}

En términos puramente estructurales, el despliegue en forma de cláusula subordinada de lo que en el original es un sintagma nominal, se interpreta como un síntoma de complejización sintáctica, y por eso, según los índices de CN y de CC, A2 es un texto levemente más complejo que A1. En lo que sigue veremos si los lectores lo perciben así también.

5.1.4. Cuando se les pregunta expresamente a los informantes por la dificultad de A1, la mayoría lo juzga entre bastante y muy difícil (con una puntuación media, del 1 al 5, de 3,95). Nadie lo considera nada o poco difícil. Cuando se les pregunta por las unidades concretas del texto, la mayoría de los lectores identifican los nombres abstractos como el elemento que más dificulta la comprensión del texto en A1 (más de medio punto por encima del siguiente elemento más complejizador, la puntuación) y señalan los conectores y los verbos en forma personal, casi a la par, como los elementos que menos lo dificultan. En A2, donde el número de sustantivos deverbales se reduce, pero siguen estando presentes, estas unidades siguen siendo valoradas como los elementos que más complejizan la comprensión de todos los propuestos, aunque ahora no lo puntúan por encima de 2 y la distancia con el siguiente elemento más complejo, las formas no personales, es de 0,4 puntos. Los conectores son, de nuevo, lo que se considera más favorable para la comprensión del texto, seguido de los verbos en forma personal. Estos resultados se ofrecen en la tabla 4:

39. Consideramos este *aunque* como conjunción coordinante, o al menos no típicamente subordinante (cf. NGLE, 2009, §§ 47.12n y 47.12ñ y Flamenco García, 1999, p. 3809 y § 59.2.3.).

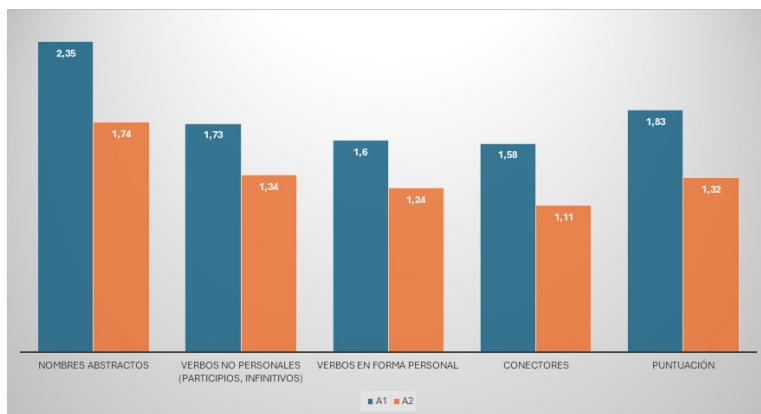


Tabla 4: Percepción del grado de dificultad de ciertas unidades morfosintácticas de los textos A1 y A2

En resumen, los informantes identifican perfectamente los sustantivos deverbales como los responsables de la complejidad textual en ambos textos, sobre todo en A1, y evalúan positivamente la presencia de conectores. Todos los elementos que valoran apuntan a una menor complejidad de A2, lo que contradice los resultados de las métricas de complejidad sintáctica presentados en la tabla 3.

5.2. *Formas no personales del verbo*

5.2.1. *Infinitivos y participios. Textos financieros.*

5.2.1.1. Algunas guías del lenguaje claro, como la GCV (§ 4.2.13), la GRJ (pp. 26-27) o la guía de la Comisión Europea, son conscientes del problema que supone para la lecturabilidad el recurso a los sustantivos deverbales y recomiendan sustituirlos por verbos, «que tienen más vida» (CEC, p. 8)⁴⁰. Sin embargo, en esta última guía, no se propone sustituir los sustantivos por formas conjugadas del verbo, sino por formas no personales.

El espíritu de esta recomendación se fundamenta en la misma idea que hemos defendido arriba: el español se maneja mejor con verbos que con nombres, pero ignora que lo que realmente facilita la comprensión es la conexión déictica del verbo con el aquí y ahora del acto de habla. Las formas no personales del verbo no son plenamente déicticas, de hecho, no son verbos plenos, pues la única categoría verbal que mantienen parcialmente es el aspecto, aunque es cierto que favorecen una predicción más similar a la de los predicados verbales, con la misma marcación

40. En Gómez Belart, Colella y Azzaro (2023, pp. 189 y 191) se menciona también tangencialmente este problema.

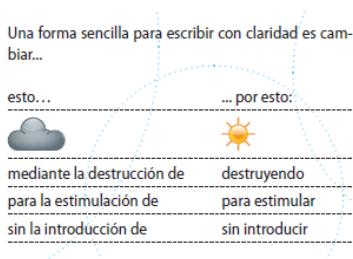


Figura 3: Comisión Europea: *Cómo escribir con claridad*, p. 8

argumental. La gran diferencia, como veremos, es que se insertan en el discurso sin la mediación de un nexo. Recuérdese que, al menos en la evaluación de los textos A, los lectores valoraban los conectores como las unidades menos «complejizadoras» para la lectura (cf. tabla 4).

En Azpiazu (2004a) se defiende que los infinitivos, participios y gerundios son también maneras de nominalizar la expresión, quizás no tan avanzadas como la nominalización deverbal, pero igualmente con efectos de desanclaje del predicado respecto al momento elocutivo. Para empezar, son formas que se insertan en las oraciones mediante lo que Lehmann (1995) y Raible (1992) denominan «coalescencia». La coalescencia se produce cuando dos o más predicados sucesivos comparten alguno de sus argumentos, de modo que este solo se expresa en el predicado principal y queda correferenciado en los demás. Para Lehmann (1995, p. 148) la coalescencia implica una mayor vinculación entre los predicados («any increase in bondedness»). Las estructuras de infinitivo, gerundio y participio son estructuras coalescentes, que, además, por ser no personales, permiten esconder al agente del evento tan bien como los predicados deverbales. Por ejemplo, podemos ver que al producir A2, ChatGPT sustituye un sustantivo deverbal del original (6) por un infinitivo (7):

- (6) A1: *salvo que el Rector (...) aprecie que* la ejecución del acto *pudiera causar prejuicios de imposible o difícil reparación*.
- (7) A2: *salvo que el Rector (...) considere que* ejecutar la decisión *podría causar daños que no se puedan reparar fácilmente*.

El resultado es una estructura un poco más verbal, pero aún nominal, donde siguen sin definirse las coordenadas deícticas del evento.

5.2.1.2. Otra forma de hacer predicaciones nominales y condensadas son los participios. No es raro encontrar en los contratos bancarios formulaciones como la siguiente:

- (8) *el Banco te entregará el dinero que tengas a tu favor en la Cuenta y tú deberás devolver todos los cheques [no utilizados]₁ [asociados]₂ a la Cuenta Vivienda, que, en una eventual reformulación, podría quedar como:*
- (9) *el Banco te entregará el dinero que tengas a tu favor en la Cuenta y tú deberás devolver todos los cheques [que no hayas utilizado]₁ y [que se asocien]₂ a la Cuenta Vivienda.*

La principal diferencia entre ambas oraciones es que en (9) se aporta información que puede ser relevante para una comprensión más precisa del texto y que no está presente en (8), como el agente de *utilizar* y la modalidad epistémica de los dos predicados.

Posiblemente, los participios (y los infinitivos) no provoquen tanto problemas de lectura como los sustantivos deverbales. De hecho, si comparamos de nuevo A1 y A2, vemos que la IA apenas ha modificado los predicados participiales presentes en A1, sino que, en general, los mantiene todos y cambia, si acaso, solo el lexema (*seleccionado* por *propuesto*). Sin embargo, para este trabajo queríamos comprobar si el empleo masivo del participio como forma nominal de predicación, con todas las relaciones de coalescencia e integración (es decir, ausencia de conectores) que conlleva, puede ser también una rémora para la lecturabilidad:

Para ello, extrajimos B1 de una póliza bancaria, de la que marcamos y numeramos las construcciones participiales:

B1. Para cumplir con lo previsto en el artículo 16.4 de la Ley 16/2011 [1], y en caso de que tu Contrato establezca que tus pagos no reducen el importe total del crédito, sino que sirven para re establecer el capital disponible del crédito concedido por el Banco [2], en las condiciones y periodos establecidos [3], las partes declaran que el Contrato no prevé una garantía de reembolso del importe total del crédito concedido [4], salvo en su caso, las garantías otorgadas por los fiadores en esta Póliza [5] u otras garantías otorgadas en Anexos a la misma [6] y, en todo caso, tu propia garantía personal,⁴¹

y le pedimos, esta vez a Gemini, que la redactara en lenguaje claro, con el siguiente resultado:

B2. Para cumplir con lo que establece el artículo 16.4 de la Ley 16/2011 [1], y si tu contrato indica que tus pagos no disminuyen el saldo

41. Anexo a una póliza de préstamo personal al consumo de Kutxabank, S.A.

total del crédito, sino que sirven para volver a poner a disposición el dinero del crédito concedido por el banco [2], bajo las condiciones y plazos acordados [3], las partes confirman que el contrato no incluye una garantía de devolución total del crédito Ø [4], salvo las garantías proporcionadas por los fiadores en este contrato [5] Ø [6] o en anexos adicionales, y siempre será válida tu propia garantía personal.

Se observa que los cambios entre estos dos textos han sido mucho menores que entre A1 y A2. En este caso, la IA ha mantenido prácticamente todos los infinitivos y participios del original; como mucho, cambia unos términos por otros (ver [3] y [5]) o directamente elimina participios del original [4] y [6]. El mayor cambio se produce en [1], donde la construcción participial (*lo previsto en...*) se desarrolla en forma de oración de relativo (*lo que establece...*). Otro añadido es el nuevo predicado verbal al final del párrafo (*y siempre será válida tu propia garantía personal*), que queda mal integrada en el conjunto del texto. Lo cierto es que, a diferencia de lo que sucedía entre A2 y A1, B2 no aporta un estilo de lengua muy diferente al de B1.

5.2.1.3. Si preguntamos a las herramientas de medición de lecturabilidad, nos dirán que, efectivamente, apenas hay diferencia entre ambos textos. Las que se basan en métricas por cantidad y longitud de palabras (*Legible* y SpR), no detectan grandes diferencias (figs. 4a y 4b). Para SpR, el balance es idéntico en los dos casos: son textos «difíciles», para estudiantes del grado 12 o 2º de Bachillerato (entre 17 y 18 años).

Legibilidad del texto		
índice	valor	dificultad
Fernández Huerta	33.44	difícil
Gutiérrez	32.13	normal
Szigriszt-Pazos	29.74	árido
INFLESZ	29.74	muy difícil
legibilidad μ	58.78	un poco difícil

Figura 4a. Índices de lecturabilidad de B1
según *Legible*

Legibilidad del texto		
índice	valor	dificultad
Fernández Huerta	37.85	difícil
Gutiérrez	32.78	normal
Szigriszt-Pazos	34.24	árido
INFLESZ	34.24	muy difícil
legibilidad μ	56.98	un poco difícil

Figura 4b. Índices de lecturabilidad de B2
según *Legible*

Curiosamente, según el índice μ de Muñoz y Muñoz, B2 sería algo más difícil de leer que B1. Por su parte, CLARA no arroja tampoco grandes diferencias entre ambos textos: 31 % de claridad para B1 y 34 % para B2:



Figura 5a. Índice de claridad de B1
según CLARA



Figura 5b. Índice de claridad de B2 según
CLARA

5.2.1.4. También las diferencias sintácticas entre los dos textos son mucho menos pronunciadas que en el caso anterior: el número de formas personales en cada uno es parecido (5 en B1 vs. 7 en B2), pero esta ratio se invierte a favor de B1 si se trata de formas no personales (infinitivos y participios):

	Verbos en forma personal	Formas no personales	Nexos
B1	5	8	4
B2	7	5	5

Tabla 5. Relación entre elementos de predicción verbal y nominal, y número de nexos oracionales en B1 y B2

En este caso, en ninguno de los dos textos hay apenas sustantivos deverbales y los que hay funcionan como predicados resultativos (*importe, reembolso*) y no como eventivos⁴².

Aún mejor que la tabla 5, la tabla 6 muestra la exacta coincidencia en la configuración sintáctica de ambos textos:

	Unidades-T	Subordinadas	Rango de subordinación	CN (Sub/U-T)	CC
B1	1	10	4	100	38
B2	1	10	4	100	38

Tabla 6. Datos de complejidad sintáctica para B1 y B2

Según los parámetros de CN y CC, ambos textos son muy complejos (superan con creces el umbral del 50 y el 5, respectivamente). En efecto, los dos presentan una gran cantidad de proposiciones subordinadas dentro de una sola unidad-T y, sobre todo, lo hacen en hasta cuatro niveles de subordinación. Sin duda, la ausencia de conectores propia de los predicados participiales e infinitivos⁴³ favorece esta incrustación sucesiva de los predicados, tal como se aprecia en (10), que pertenece a B1:

42. Sobre las diferencias sintácticas y semánticas entre sustantivos deverbales resultativos y eventivos en español, cf. Picallo (1999).

43. Consideramos que una forma no personal (infinitivo o participio, en este caso) configura una cláusula subordinada si dicha forma es el núcleo de un predicado con expansión hacia la derecha a través de argumentos o adjuntos.

(10) *en caso de*

[que *tu Contrato* establezca
 2 [que *tus pagos no* reducen *el importe total del crédito*, sino que sirven para
 3 [re establecer *el capital disponible del crédito*
 4 [concedido por *el Banco*]]]]₂₁

En definitiva, hay que pensar que para la inteligencia artificial B1 no es un texto que merezca mucha manipulación, es decir, los participios y el tipo de construcción sintáctica al que dan lugar no cuentan como estructuras especialmente ajenas al lenguaje claro en español.

5.2.1.5. Si recurrimos a los propios lectores, vemos que estos textos los juzgan más benévolamente que los anteriores: para empezar, no consideran B1 tan difícil como A1: el 46 % lo considera bastante difícil, pero solo el 14 % dice que es muy difícil. Su puntuación media es de 3,6 sobre 5⁴⁴. La diferencia en la lectura de ambos textos también apunta en esa misma dirección: de media tardan casi 10 segundos menos en leer B1 que A1, con prácticamente el mismo número de palabras (A1 = 94 palabras; B1 = 99 palabras)⁴⁵. Una representación gráfica de estas diferencias, en la que se incluye también C1, puede verse más abajo (tablas 10 y 11).

Respecto a los elementos que les producen mayor dificultad, en B1 aparecen de nuevo los sustantivos deverbales en primera posición, aunque con una media inferior a la que aparecía en el texto A1 (2,14 % vs. 2,35 de A1), seguido de la puntuación. Los demás elementos se ordenan de manera levemente diferente a A1: en B1 los verbos en forma personal superan a los conectores como el elemento de menor complejidad, aunque lo cierto es que los porcentajes son prácticamente los mismos que en A1.

En la tabla 7 vemos representadas las evaluaciones de todos los elementos en los cuatro textos. Como se aprecia, las puntuaciones repiten casi exactamente el mismo patrón para los textos originales (líneas superiores) que para los textos facilitados (líneas inferiores). Los únicos elementos que subrayan la conciencia de una diferencia entre ambos textos son los sustantivos deverbales en A1, como elementos especialmente complejos de asimilar, y, en menor medida, los conectores y la puntuación en A2,

44. Es interesante, con todo, que a pesar de esta supuesta menor dificultad, los lectores lo entienden peor que A1: el índice de acierto de la pregunta de comprensión es del 57 %, frente al 69 % de acierto de A1.

45. Otro índice ilustrativo de la diferencia en la manera de procesar un texto y otro lo tenemos en el dato de la desviación estándar. Con este índice sabemos que hay grandes diferencias en el tiempo de lectura entre los informantes. Este valor es muy alto (36,43) en A1 y algo inferior, aunque también alto (30,07) en B1.

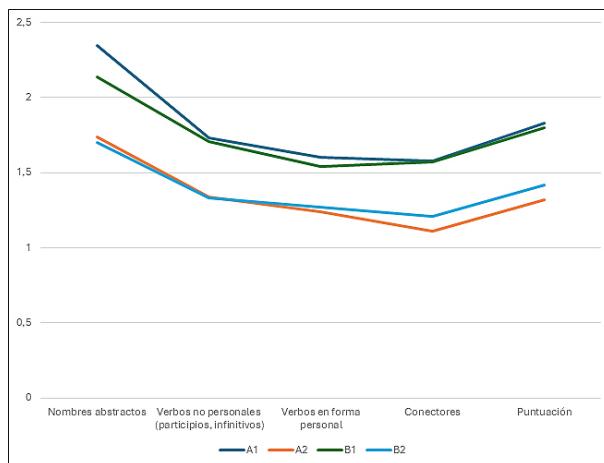


Tabla 7. Percepción del grado de dificultad de ciertas unidades morfosintácticas de los textos A1 - A2 y B1 - B2 por los informantes

como lo contrario. Es también interesante que, a pesar de que sí hay conectores en B₂, los lectores no los consideran tan favorables para la lecturabilidad como en A₂. Una posible explicación puede ser que en este caso quedan diluidos en los incisos y en los sucesivos niveles de subordinación. Quizá los lectores competentes de español procesan aceptablemente bien este nivel de complejidad sintáctica, pero es posible que echen de menos algún elemento de conexión oracional más.

5.2.2. *Gerundios. Textos judiciales.*

5.2.2.1. La misma coalescencia y condensación con pérdida de marcas gramaticales que permiten el infinitivo y el participio es propia también de la tercera forma no personal del verbo, el gerundio, en función predicativa. Sin embargo, las diferencias diatéticas y aspectuales entre los participios (pasivos, resultativos) y los gerundios (activos, progresivos) hacen que el uso de uno u otro afecte de manera muy distinta a la dinamicidad y, con ello, a la lecturabilidad del texto. Hay que señalar que el gerundio es, de las tres formas no personales del verbo, la que peor se ha integrado siempre en el sistema (y en el estilo) de la lengua. Todos los manuales consultados que se ocupan específicamente del español (GGV, MGA, MAM; cf. también Gómez Belart, Colella y Azzaro, 2023, pp. 196-197) dedican algún apartado a censurar su uso. Como veremos, los problemas que presenta son también de naturaleza algo diferente a los analizados hasta ahora.

El siguiente texto, C1, que es la parte del relato de los hechos de una sentencia judicial, es un ejemplo paradigmático de uso espurio del gerundio (se marcan y numeran ahora solo estas construcciones):

C1. Tras acceder al referido lugar, el acusado Aureliano, portando una pistola negra con apariencia de arma de fuego y con la cara al descubierto [1] y el acusado Eladio, portando un arma blanca tipo puñal de 25-30 centímetros de hoja [2], se dirigieron a dos empleadas de dicha sucursal, Paulina y Violeta y al director de la misma Nicolas, solicitándole a Violeta el dinero del cajero automático [3], realizando dos expediciones de 300 euros cada una de ellas [4], apoderándose de 600 euros en billetes y 67,78 euros en metálico [5] así como los terminales móviles de los empleados, metiendo a los tres empleados en el despacho del director [6] rociándoles con un spray lacrimógeno [7] para facilitar su huida y no ser perseguidos.⁴⁶

A diferencia de los textos anteriores, meramente expositivos y atemporales, este nos presenta una narración de eventos del pasado. Sin embargo, es una narración anómala. Toda ella se estructura en torno a un único predicado en forma personal (*se dirigieron*) al que se le adjunta, en forma de subordinadas sin un solo conector ni otro verbo conjugado, una sucesión de eventos que no se anclan en el tiempo, pues solo presentan temporalidad interna. La anomalía del texto proviene de la decisión de encargar al gerundio, una forma verbal no deictica y de aspecto imperfectivo, es decir, discursivamente descriptiva, la labor de marcar todos los puntos prominentes de la narración. El efecto es extraño, porque los gerundios hacen hincapié en el proceso y no en la terminación de los eventos, de manera que no son formas apropiadas para señalar estados de cosas sucesivos que permitan progresar una narración⁴⁷. Empleado de esta forma, el gerundio genera una especie de «masa» de eventos que se superponen pero que no parecen avanzar porque no se anclan en el tiempo, y cuyos agentes son siempre los mismos dos individuos a los que se hace referencia en el único verbo en forma personal que ejerce de pivote de toda la narración. Finalmente, es el lector, basándose en el orden en el que se

46. Fragmento del Documento TOL5.776.076 del *Corpus RLD: el corpus de uso de recursos lingüísticos en el derecho español*, Universidad Complutense (<https://docta.ucm.es/entities/publication/3d6203cb-21ec-4f19-a053-7bdd1282110d>).

47. Este es precisamente el problema del «gerundio de posterioridad», que se censuran en las gramáticas y también en las guías del lenguaje claro (ver, por ejemplo, MAM, p. 54; GRJ, p. 25; Gómez Belart, Colella y Azzaro, 2023, p. 197) y del que hay varios ejemplos en este texto.

presentan los hechos, en el contexto y en su propio sentido común, el que debe reconocer aquí una narración.

En definitiva, la complejidad de la lectura de este fragmento es de naturaleza muy diferente a la de los textos anteriores: en este caso el problema no es tanto la abstracción y la condensación del contenido (aunque la coalescencia y la falta de deixis puede complicar la inteligibilidad)⁴⁸, como que, a pesar de toda la información que se ofrece en torno a un único predicado pleno, faltan las marcas narrativas al uso. Como veremos, los problemas de lecturabilidad de este fragmento provienen no tanto de su complejidad textual, que es alta, pero no parece molestar a la lectura, como de su estilo narrativo o, si se quiere, de su ausencia del mismo.

Al pedirle a ChatGPT que facilitara el texto, el resultado fue el siguiente:

C2. Aureliano portaba una pistola negra con apariencia de arma de fuego y llevaba la cara descubierta [1], mientras que Eladio portaba un arma blanca tipo puñal de entre 25 y 30 cm de hoja [2]. Una vez en el lugar, ambos se dirigieron a las empleadas del banco, Paulina y Violeta, así como al director, Nicolás, exigiéndoles el dinero del cajero automático [3]. Obligaron a Violeta a realizar dos extracciones de 300 euros cada una [4], apoderándose de un total de 600 euros en billetes y 67,78 euros en monedas [5], además de los teléfonos móviles de los empleados. Para facilitar su huida y evitar ser perseguidos, encerraron a los tres empleados en el despacho del director [6] y los rociaron con un spray lacrimógeno [7].

En esta ocasión la modificación ha sido bastante radical. Lo primero que se aprecia es que el texto gana en textura discursiva, precisamente porque se ha pasado de una a siete formas personales del verbo de dos tipos tempoaspectuales diferentes, lo que permite diferenciar sendas funciones textuales: descripción ([1] *Aureliano portaba una pistola negra*; [2] *Eladio portaba un arma blanca*) y narración ([4] *obligaron*, [6] *encerraron*, [7] *rociaron*). Repárese en que aún se mantienen algunos gerundios de posterioridad innecesarios ([3] *exigiéndoles*; [5] *apoderándose*), pero en general el episodio se ancla en unas coordinadas temporales claras que permiten seguir la sucesión lógica de los eventos.

48. En este mismo sentido, es muy interesante que en la GRJ, destinada precisamente a este tipo de textos, se advierte contra lo que se denominan «gerundios ilativos» y se hace una precisión que apunta en la misma dirección señalada aquí: «cuando se acumulan formas de gerundio, suele resultar difícil identificar el verbo flexionado con el que hay que relacionarlas, y cuesta entender también, clara y rápidamente, quién lleva a cabo la acción y cuándo la hace» (p. 25).

5.2.2.2. Cuando preguntamos a los programas de lecturabilidad, tanto *Legible* como SpR consideran C1 entre «difícil» y «muy difícil». Según estos valores, este texto es el que peor lecturabilidad tiene de todos los analizados. Por el contrario, C2 se valora como el más lecturable de todos⁴⁹. Para SpR, C1 es difícil y propio del nivel escolar 12 (17-18 años, 2º de Bachillerato), mientras que C2 es «ligeramente difícil» y propio del nivel 10 (15-16 años, 4º de la ESO).

Legibilidad del texto		
índice	valor	dificultad
Fernández Huerta	17.9	muy difícil
Gutiérrez	27.97	normal
Szigriszt-Pazos	14.31	muy difícil
INFLESZ	14.31	muy difícil
legibilidad μ	49.05	difícil

Figura 6a. Índices de lecturabilidad de C1
según *Legible*

Legibilidad del texto		
índice	valor	dificultad
Fernández Huerta	54.97	algo difícil
Gutiérrez	40.68	normal
Szigriszt-Pazos	50.72	normal
INFLESZ	50.72	algo difícil
legibilidad μ	53.18	un poco difícil

Figura 6b. Índices de lecturabilidad de C2
según *Legible*

Sin embargo, para CLARA, la probabilidad de que C1 sea un texto claro es de un 58 %, es decir, bastante alto, si lo comparamos con los porcentajes que arroja para el resto de los textos originales, que no superaban el 25 %, mientras que en C2 es del 74 %. Podríamos decir que para CLARA, C2 es un texto narrativo casi típico en español. En cualquier caso, es importante que en ambos casos el porcentaje de claridad supera el 50 %, lo que no sucede con ninguno de los textos facilitados anteriores.



Es **poco probable** que el texto sea claro

Figura 7a. Índice de claridad de C1 según
CLARA



Es **bastante probable** que el texto sea claro

Figura 7b. Índice de claridad de C2 según
CLARA

Esta discrepancia con los otros instrumentos de medición no sorprende si recordamos que CLARA es la única herramienta que tiene en cuenta aspectos morfosintácticos y discursivos que trascienden el mero cómputo de palabras y de sílabas.

5.2.2.3. En cuanto a la complejidad sintáctica, la tabla 8 muestra la gran discrepancia en el empleo de las formas personales y no personales, así como en el uso de conectores, entre C1 y C2. Estos datos se amplifican en la tabla 9, donde vemos que

49. Aunque C2 supera el 50 % en el test μ de Muñoz y Muñoz, da un valor algo menor de lecturabilidad que los dos textos financieros, tanto B1 como B2. En realidad, según este test, B1 es el texto más lecturable de los seis analizados. Presentamos este dato como mera curiosidad y para mostrar cómo las métricas basadas únicamente en número de palabras y número de letras por palabra no pueden servir como medidas razonables de complejidad textual en español (posiblemente, en ninguna lengua).

C₁ consta de una sola unidad-T, cuyo núcleo es el único verbo en forma personal de todo el texto, a la que acompañan hasta diez cláusulas subordinadas sin nexo ([1] a [7], más otras cláusulas de infinitivo). Aunque estas subordinadas son en su mayoría constituyentes directos del predicado principal, es decir, se encuentran en el primer nivel de subordinación, su abundancia en el texto y la desproporción respecto a la unidad-T dispara el índice de complejidad sintáctica, numérica y cualitativa, de este texto. Por su parte, los índices se normalizan en *C₂*, donde predominan las formas personales sobre las no personales, hay seis unidades-T, pero la mitad de subordinadas que en *C₁*, además de cuatro conectores. Como se aprecia, en este texto los índices están dentro de los parámetros que Salvador Mata (1985) consideraba propios de un texto claro y coherente (CN ampliamente por debajo de 50 y CC por debajo de 5):

	Verbos en forma personal	Formas no personales	Nexos
<i>C₁</i>	1	10	1
<i>C₂</i>	7	5	4

Tabla 8. Relación entre elementos de predicción verbal y nominal, y número de nexos oracionales en *C₁* y *C₂*

	Unidades-T	Subordinadas	Rango de subordinación	CN (Sub/U-T)	CC
<i>C₁</i>	1	10	2	100	30
<i>C₂</i>	6	5	2	8	4.3

Tabla 9. Datos de complejidad sintáctica para *C₁* y *C₂*

Los índices de complejidad sintáctica de *C₁* representan una anomalía de tipo textual: lo que podría haberse presentado como diferentes unidades-T, una para cada evento de la narración, son cláusulas subordinadas que complementan paralelamente a un único y mismo predicado. La pregunta ahora es hasta qué punto esta extraña configuración textual supone un problema para su lecturabilidad según los propios lectores.

5.2.2.4. Por una parte, *C₁* es sin duda el texto que más rápidamente se lee de media (a pesar de tener el mayor número de palabras, 114)⁵⁰. Por la otra, en la encuesta

50. También el dato de la desviación estándar (DE) es significativo aquí. Frente a 36,43 de *A₁* y 30,07 de *B₁*, en *C₁* la DE es de 21,18, notablemente menor. La DE se reduce bastante en los textos facilitados, y, aunque sigue siendo alta, las diferencias entre ellos no son tan pronunciadas como en los originales: *A₂* = 18,39; *B₂* = 15,65; *C₂* = 13,89.

ningún informante clasifica C1 ni como bastante ni como muy difícil, sino que lo puntúan en los tres valores inferiores, con una media de un 2,13. Es, sin duda, el texto original que más fácil les resulta de los tres, a bastante diferencia de los otros dos.

En las tablas 10 (a y b) y 11 (a y b) se aprecian más claramente las diferencias en la recepción de los tres textos:



Tabla 10a: Tiempo medio de lectura de los textos originales

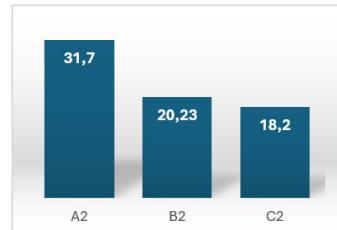


Tabla 10b: Tiempo medio de lectura de los textos facilitados



Tabla 11a: Valoración media de dificultad por parte de los lectores de los textos originales



Tabla 11b: Valoración media de dificultad por parte de los lectores de los textos facilitados

Las tablas 10 nos muestran claramente que los textos con más sustantivos deverbales son siempre más lentos de procesar que los que tienen formas no personales del verbo, sean estos gerundios o participios. También vemos que la diferencia en el tiempo de lectura entre B y C no es tan pronunciada, sobre todo en los textos facilitados, pero en estos mismos textos, la diferencia con A2, que aún mantiene varias construcciones nominalizantes del original, se hace significativamente grande⁵¹. En cuanto a la propia percepción de los informantes (tablas 11), se aprecia claramente que los textos C se perciben siempre como de fácil lecturabilidad, frente

51. El valor p que arrojan los datos de la tabla 10a (textos originales) es de 0,10 (las diferencias entre los textos son poco significativas), mientras que el de la tabla 10b (textos facilitados) es de 0,002 (esas mismas diferencias son muy significativas estadísticamente).

a A y B, que consideran difíciles o muy difíciles⁵². Curiosamente, incluso, el texto facilitado B₂ les resulta algo más complejo que A₂, lo cual posiblemente tenga que ver con el hecho de que, como se ha demostrado, es también el texto que menos modificaciones desde el original ha sufrido de los tres. Los índices de A₂ y B₂ que arroja *Legible*, presentados en las figuras 1b y 4b, respectivamente, apuntan en esta misma dirección. Con todo, los informantes desmienten los resultados de *Legible* respecto a C₁, corroboran casi todos los de CLARA y, de nuevo, demuestran que los índices de complejidad sintáctica no se corresponden con su propia concepción subjetiva de lo que es un texto difícil de leer.

5.2.2.5. La poca dificultad que encuentran los lectores en C se traslada a su valoración de los elementos concretos, pues ninguno de ellos supera el umbral del 2 en ninguna de las dos versiones. En el caso de C₂, las puntuaciones se mueven en torno al grado mínimo y ninguna de ellas destaca sobre el resto⁵³. En C₁, el elemento que, según los lectores, dificulta más la lectura es ahora la puntuación, seguido, a bastante distancia, de los gerundios.

En la tabla 12 se comparan los resultados obtenidos en esta pregunta para C₁ y C₂ con los del resto de los textos, es decir, se replica la tabla 7 y se añaden los resultados de estos dos últimos textos:

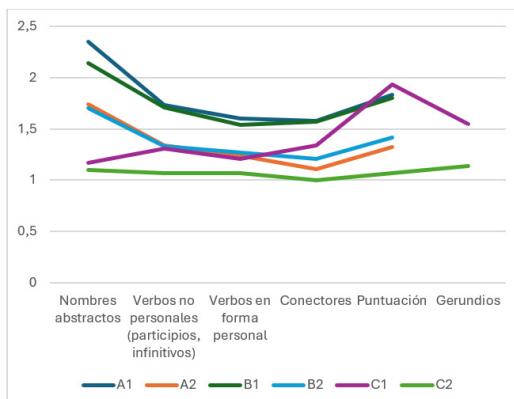


Tabla 12: Percepción del grado de dificultad de ciertas unidades morfosintácticas de todos los textos de la encuesta, según los informantes

52. Sin duda, el hecho de que los textos C presenten una narración y los otros sean expositivos e incorporen léxico jurídico-administrativo juega también un papel en la distinta percepción de su dificultad.

53. Según el análisis ANOVA de un factor, la diferencia de complejidad percibida entre los elementos es significativa en C₁ (como en el resto de los textos analizados) ($p < 0,01$), pero no lo es en C₂ ($p > 0,05$).

Del gráfico se desprende que los informantes calibran la dificultad de los textos C de forma muy diferente a la de los textos A y B, mucho más homogéneos entre sí. Los nombres abstractos han dejado de ser un problema; obviamente, porque apenas los hay aquí, pero la puntuación se convierte ahora en un obstáculo serio para la comprensión⁵⁴. Este gráfico incluye, además, la valoración de los gerundios, que no se testaba en los otros dos textos y que, como era esperable, se puntúa también como un elemento distorsionador.

Creemos que estos resultados pueden interpretarse en la línea señalada antes: C1 es una sucesión de eventos yuxtapuestos, con apariencia de narración pero sujeta a una forma antinarrativa. También podría interpretarse en los términos de integración/agregación de Raible (1992): se desarrolla una narración de eventos sucesivos y yuxtapuestos, pero se presenta bajo la forma de predicados semiintegrados a través de la coalescencia de los sujetos. No representa, por tanto, ningún sistema de junción bien identificable en la escala de la tabla 1. En general, podríamos pensar que los lectores son capaces de leer sin grandes dificultades C1, pero esta forma de contar las cosas no encaja en sus expectativas idiomáticas. ChatGPT parece darse cuenta de ello también: C2 se acerca mucho a un texto narrativo «normal» en español, donde integración y agregación se combinan equilibradamente y ningún elemento destaca por ofrecer problemas especiales a la lectura⁵⁵.

6. CONCLUSIONES

6.1. Los resultados de la encuesta apoyan la primera hipótesis, propuesta en § 4.1: el empleo de categorías nominales con fines predicativos, es decir, el estilo nominal, parece complicar la lecturabilidad de los textos en español, al menos así lo perciben tanto los hablantes como los programas de IA generativa.

Los elementos nominales que más entorpecen la lectura son, sin duda, los sustantivos deverbales empleados predicativamente. Las formas no personales del verbo son también predicados pseudonominales que contribuyen al estilo nominal. Complican algo la lectura, pero lo hacen de maneras diferentes:

a) Los infinitivos están bastante integrados como estructuras subordinadas sustantivas. Implican más nominalización que las cláusulas subordinadas con verbo flexionado, pues predicen sin nexos, pero menos que los sustantivos deverbales.

54. Posiblemente, lo que condenan los informantes es que todo el texto sea una secuencia larga de eventos no interrumpida por ningún punto ni orientada por ningún conector.

55. Cf. nota 54.

b) Los participios se procesan sin mucho problema por su proximidad funcional con los adjetivos; incluso la IA generativa los pasa por alto cuando se le encarga la labor de facilitación, a diferencia de lo que hace con los sustantivos deverbales. Tampoco se citan apenas en los manuales de lenguaje claro consultados como elementos de complejización textual. Sin embargo, hemos visto que favorecen un alto nivel de complejidad sintáctica y que pueden crear ambigüedad semántica, al esconder mucha de la información deíctica propia del verbo.

c) Los gerundios no dificultan especialmente la lectura, pero provocan extrañeza en contextos narrativos porque atentan contra el estilo de la narración en español. En el contexto de C1 su uso aporta, quizás, más problemas de aceptabilidad que de inteligibilidad. El texto se entiende, pero es «mal español».

6.2. Como se ha argumentado, la predicación por medio de elementos nominales y no verbales implica complejidad a nivel semántico (abstracción del contenido) y sintáctico: las estructuras se condensan en menos palabras, pero sus relaciones se vuelven opacas. Esto significa que el nivel de lecturabilidad de un texto no puede medirse solo con criterios como la cantidad o la longitud de las palabras, como se hace en *Legible* o *Spanish readability*. Tampoco con las fórmulas de medición de complejidad sintáctica al uso, porque, como hemos podido comprobar, una cierta complejidad sintáctica no implica necesariamente, en un adulto con una experiencia lectora estándar⁵⁶, problemas de lecturabilidad.

Comparemos, por ejemplo, el siguiente gráfico, que aúna los resultados de las pruebas de complejidad numérica y cualitativa en todos los textos, con las tablas 10 (de tiempo de lectura) y 11 (de valoración del grado de dificultad) anteriores (§ 5.2.2.4).

La tabla 13 pone de manifiesto que B1, B2 y C1 son los textos que *a priori* presentan mayor complejidad sintáctica, pero también que la complejidad de A1 es pequeña, inferior incluso a la de A2. Recordemos, sin embargo, que según los informantes (tablas 11a y b), A1 es sin duda el texto más difícil de leer de todos,

56. Es obvio que la validez de esta afirmación implica definir qué entendemos por «experiencia lectora estándar». Podría aducirse que, si esta experiencia es la propia de los estudiantes universitarios, grupo al que pertenecen casi todos los informantes de este estudio, muchos usuarios de la lengua carecerán de ella. Sin embargo, no creo que esto sea cierto. En mi opinión, en una sociedad con un nivel de escolarización amplio y en la que los individuos conviven cotidianamente con un mundo gráfico y textual diverso, la «experiencia lectora estándar» es la propia de la mayoría de los adultos. Ello permite una cierta homogeneización de la competencia lectora básica, sin que, desgraciadamente, se garantice que los individuos con estudios superiores sean necesariamente lectores más competentes que el resto. No obstante, es cierto que todo esto debería analizarse con más precisión, si queremos que este concepto nos sirva para medir la validez de los juicios de nuestros informantes.

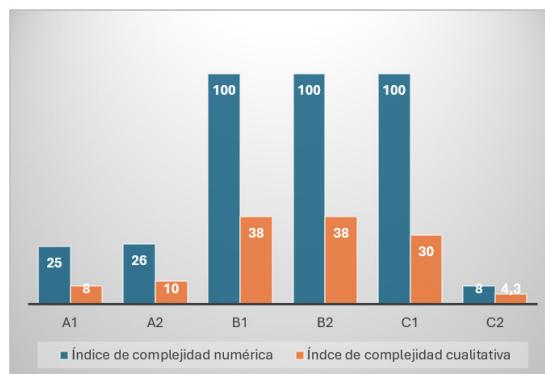


Tabla 13. Complejidad numérica y complejidad cualitativa de los seis textos de la encuesta

que C1 es poco complicado y que los textos B se encuentran en la encuesta en un término medio de dificultad, algo que no se reflejan en absoluto en este gráfico.

Todo esto nos hace pensar que los resultados también apoyan la segunda hipótesis de § 4.1.: la llamada «complejidad sintáctica» no es necesariamente un gran problema de lecturabilidad para un usuario con un uso competente pero no extremadamente sofisticado del español. La tendencia al estilo verbal de la lengua hace que un cierto grado de complejidad sintáctica sea perfectamente asumible por los hablantes. Por lo tanto, además del número y la longitud de las palabras y las oraciones, postulamos que hay que empezar a tomar en consideración y a calibrar con algo más de utilería otros factores de tipo morfosintáctico, como algunos de los señalados en este trabajo: presencia o no en el discurso de formas personales del verbo, de sustantivos predicativos, de cierto tipo de conectores, uso sintáctico de la puntuación, etc., y valorar, no solo cuantitativa, sino también cualitativamente, la interacción de estos elementos en el texto.

6.3. En definitiva, quizá no sea posible definir exactamente en qué consiste el lenguaje claro en español, pero tal vez sí pueda hacerse un «retrato robot» al que debe aspirar para funcionar como un código que permita a la mayoría de los hablantes de una lengua acceder a los textos que lee: propongo que sea uno que respete la tendencia natural del español a un estilo verbal más que nominal, es decir, que tienda a predicar con formas personales del verbo más que con las no personales o con nominalizaciones y que se encuentre en un punto intermedio entre la máxima agregación y la máxima integración sintáctica, esto es, que tienda a concatenar las ideas a través de marcadores explícitos y no mediante signos de puntuación o marcas gramaticales excesivamente abstractas.

Ahora bien, hay varios peligros evidentes de manejar esquematizaciones como esta. En primer lugar, está el riesgo de tomarlas como normas que deben imperar en todo texto y en todo lugar. Obviamente, hay razones poderosas para recurrir a veces a la predicación nominal; la cuestión es no convertir esta predicación en la norma, es decir, en un «estilo» de hacer las cosas. En segundo lugar, hay que procurar mantener separados los conceptos de «claridad» y «sencillez» (cf. nota 7). Cuando esto no se hace, se puede llegar al absurdo del MGA, donde se recomienda prescindir de palabras gramaticales claves como los pronombres y algunos conectores (p. 21), los adjetivos y adverbios o incluso las perifrasis verbales, como el presente continuo o el futuro próximo, porque aportan «información accesoria» y aumentan, se supone que innecesariamente, el número de palabras del texto (p. 35; cf. también nota 13). Aquí se está confundiendo «claridad» lingüística con «minimalismo» lingüístico. El minimalismo implica la máxima sencillez en el nivel formal, pero esta no va de la mano de la sencillez en el nivel del significado o del sentido, porque hacerse inteligible en una lengua no es nunca un proceso «sencillo» en sí mismo y conseguir un mensaje claro a veces requiere añadir y no quitar palabras o estructuras.

En tercer lugar, un «retrato robot» del lenguaje claro podría ser lo suficientemente general como para encajar en el perfil de muchas lenguas diferentes. El riesgo consiste en aplicarlo sin más de una a otra, sin atender antes a la idiosincrasia morfosintáctica de cada lengua. En otras palabras, no deberían aplicarse a una lengua determinada normas de lenguaje claro dictadas para otra lengua, si no han sido convenientemente testadas antes, como se ha tratado de hacer aquí. Por ejemplo, podemos sospechar que las normas para un lenguaje claro en inglés no tienen por qué ser válidas para el español si conocemos la función que tiene en cada una de ellas el gerundio y, en general, todas las formas no personales del verbo. Para entender el modo en que se aplica el lenguaje claro en cada lengua hay que conocer bien sus diferencias de estilo, para lo cual suelen ser útiles los estudios contrastivos o traductológicos (cf. Azpiazu 2004a, pp. 22-23; 2006, pp. 73-75; 2013, p. 44).

Por último, el cuarto error, vinculado con los anteriores, consiste en pensar que el lenguaje claro es equivalente al uso de la lengua de un hablante extranjero no totalmente competente, aproximadamente con un nivel B2. Pero poner los conocimientos de la lengua de un hablante no nativo como modelo de lo que es un lenguaje claro es no haber entendido cómo funcionan las lenguas; desde luego, es no entender cómo se aprenden cuando son la L2, pero tampoco cómo se adquieren y se manejan cotidianamente cuando son la L1. Si hay algo que singulariza el conocimiento de la lengua nativa frente a la extranjera es, más que la corrección

gramatical, su heterogeneidad. Un hablante nativo medio maneja con más o menos naturalidad lo que Coseriu (1955) denominaba la «arquitectura de la lengua», y eso le permite conocer y usar en momentos precisos términos, formas y estructuras de diferentes variedades y estilos; términos y estructuras a los que un extranjero, por muy bien que conozca la lengua, raramente tiene acceso completamente. No olvidemos que un hablante nativo, por muchas dificultades que tenga para leer ciertos textos, *vive* en su lengua, no solo en el sentido pragmático (hace su vida en esa lengua), sino, sobre todo en el sentido «heideggeriano»: su lengua *es* su casa. Y esta casa tiene un estilo, unas «normas de habitabilidad» que para él son familiares, pero para un extranjero, por muy cerca que se encuentre su lengua de la que está aprendiendo o por muy bien que la conozca, no tienen por qué serlo. Así, un estudiante de ELE de nivel B2 puede entender mejor un texto con palabras comunes y no demasiado largas, oraciones simples y poca subordinación, pero para un hablante nativo, incluso uno con una formación escasa, un texto con cláusulas en diferentes grados de subordinación puede ser preferible a otro con menos palabras y un alto nivel de integración sintáctica.

6.4. Retomamos así lo que se apuntó en § 1. acerca de la afinidad entre los conceptos de «estilo de lengua» y «lenguaje claro»: son conceptos idiomáticos, sujetos a la idiosincrasia de cada lengua y difícilmente traducibles de una a otra. Según como yo los entiendo, el «lenguaje claro» es un término funcional: designa un desiderátum de inteligibilidad. El «estilo de lengua» es una forma de expresión que, en última instancia y, entre otras cosas, se vincula con lo que hacen los hablantes cuando quieren ser inteligibles. Ambos conceptos se mueven en un ámbito similar, aunque para fijar sus límites es preciso tener en cuenta la complejidad del diasistema de cada lengua⁵⁷.

6.5. Finalmente, unas palabras sobre el papel de la inteligencia artificial en la tarea de simplificación o clarificación sintáctica de los textos. Lo que hemos podido comprobar es que, al menos cuando los textos son breves, el resultado que ofrece es bastante aceptable. Su rendimiento es bueno con los casos más extremos: aunque no los evita del todo, reconoce que demasiados sustantivos deverbales no favorecen la inteligibilidad y, obviamente, que los gerundios no son formas sensatas de narrar en español. Sin embargo, en los niveles intermedios de complejidad, cuando se trata de reconocer la adecuación de otras formas de nominalización más sutiles y más integradas en el sistema de la lengua, como los infinitivos y, sobre todo, los

57. Cf. los problemas para identificar este ámbito que se mencionan en la nota anterior.

participios, no parece (todavía) ser capaz de identificar tan bien los problemas de inteligibilidad que provocan estas formas. Es en el reconocimiento de estos aspectos donde puede merecer la pena seguir trabajando con ella.

AGRADECIMIENTOS

Quiero agradecer su ayuda a Olga Ivanova, profesora del departamento de Lengua Española de la Universidad de Salamanca, por haberme introducido en el empleo de la herramienta *Qualtrics* y haberme dado pautas para trabajar con los resultados obtenidos con ella. También a Alberto Cantera, de la Universidad Libre de Berlín, por sus atinadas observaciones a la primera versión de este texto, y a los revisores anónimos del artículo, por las correcciones e indicaciones que me han hecho y que, sin duda, han contribuido a mejorarlo sustancialmente.

FINANCIACIÓN

Este trabajo forma parte del proyecto de investigación, financiado por la Junta de Castilla y León con fondos FEDER–Unión Europea, *COMPLETEXT: Complejidad textual y lecturabilidad. Estudio aplicado a la IA generativa y la didáctica de las lenguas* (ref. SA067P24). Sus IP son Susana Azpiazu Torres y Javier De Santiago Guervós (Universidad de Salamanca).

REFERENCIAS

Guías del lenguaje claro

- Ayuntamiento de Madrid - Universidad Rey Juan Carlos (s.f.). *Manual del lenguaje administrativo del Ayuntamiento de Madrid*. <https://www.madrid.es/UnidadesDescentralizadas/Calidad/Publicaciones/Documentaciontecnica/ficherodocte/ManualLA.pdf> [MAM] [Último acceso: 10/07/2025]
- European Commission. (2011). *Cómo escribir con claridad*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2782/259493> [CEC] [Último acceso: 10/07/2025]
- Generalitat Valenciana (2023). *Guía de lenguaje claro de la Generalitat Valenciana. Por un lenguaje administrativo al servicio de las personas*. <https://www.gva.es/documents/89863/89923/Gu%C3%A3da+de+lenguaje+claro+de+la+Generalitat+Valenciana.pdf/4de1f4c5-60a8-e884-47ed-dab5969be611?t=1706774826759> [GGV] [Último acceso: 10/07/2025]
- Gobierno de Aragón (s.f.). *Manual de estilo de lenguaje claro. Criterios para la redacción de textos claros en internet*. <https://www.aragon.es/documents/20127/94900387/Manual+de+estilo+de+lenguaje+claro-Gobierno+de+Aragon.pdf/15a93b64-cac4-93ca-9146-b66747f5adb3?t=1683022477318> [MGA] [Último acceso: 10/07/2025]

Ministerio de Justicia del Gobierno de España (2013). *Guía de redacción judicial clara. El derecho de la ciudadanía a entender: claves para redactar documentos judiciales eficaces*. <https://www.mjusticia.gob.es/es/AreaTematica/DocumentacionPublicaciones/InstListDownload/Gu%C3%A9rda%20redacci%C3%B3n%20judicial%20clara.pdf> [GRJ] [Último acceso: 10/07/2025]

Estudios lingüísticos

- Azpiazu, S. (2000). La predicación nominal: Infinitivos y *nomina actionis*. In J. Borrego, J. Fernández González, L. Santos Río y R. Senabre Sempere (Eds.), *Cuestiones de actualidad en lengua española* (pp. 13-22). Ediciones Universidad de Salamanca-Instituto Caro y Cuervo.
- Azpiazu, S. (2004a). *Las estrategias de nominalización. Estudio contrastivo del estilo nominal*. Peter Lang.
- Azpiazu, S. (2004b). ¿Qué es una lengua nominalizante? In M. Villayandre Llamazares (Coord.), *Actas del V Congreso de Lingüística General* (vol I., pp. 295-305). Arco/Libros.
- Azpiazu, S. (2006). Stylistic-contrastive analysis of the nominality/verbality in the languages. In C. Mourón Figueroa y T.I. Moralejo Gárate (Eds.), *Studies in Contrastive Linguistics* (pp. 69-77). Publicacíons Universidade de Santiago de Compostela.
- Azpiazu, S. (2013). El *estilo de lengua*. Antecedentes y fundamentos metodológicos. *Dicenda*, 31, 27-55. doi: https://doi.org/10.5209/rev_DICE.2013.v31.43629.
- Bień, J. (2013). *El estilo nominal en español y en polaco*. Wydawnictwo Muzyczne Polihymnia.
- Borreguero, M. (2006). Naturaleza y función de los encapsuladores en los textos informativamente densos (la noticia periodística). *Cuadernos de Filología Italiana*, 13, 73-95.
- Borreguero, M. (2018). Los encapsuladores anafóricos: Una propuesta de clasificación. *Caplletra Revista Internacional de Filología*, 64, 179-203. doi: <https://doi.org/10.7203/Caplletra.64.11380>
- Camacho Muñoz, R. (2023). El lenguaje claro, una llave para abrir la transparencia. *Transparència i Govern Obert. Sección: Sala de ponents*. [En línea]. <https://governobert.diba.cat/news/2023/03/16/llenguatge-clar-clau-per-abrir-transparencia>
- Chomsky, N. (1970). Remarks on Nominalization. In R. A. Jacobs y P. S. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (184-221). Ginn and Company.
- Coseriu, E. (1955). La geografía lingüística. *Revista de la Facultad de Humanidades y Ciencias (Montevideo)* 14, 29-69.
- Crespo Allende, N., Alfaro Faccio, P. y Góngora Costa, B. (2011). La medición de la sintaxis: evolución de un concepto. *Onomázein*, 24, 155-172. doi: <https://doi.org/10.7764/onomazein.24.07>
- Enkvist, N. E. (1987). Estilística, lingüística del texto y composición. In E. Bernárdez Sanchís (Coord.), *Lingüística del texto* (pp. 131-150). Arco/Libros.

- Escandell, M.V. (1995). *Los complementos del nombre*. Arco/Libros.
- Fernández Pérez, M. (1984): El carácter de la ciencia lingüística. *Verba*, 11, 129-156.
<http://hdl.handle.net/10347/4929>
- Flamenco García, L. (1999). Las construcciones concesivas y adversativas. In I. Bosque y V. Demonte (Eds.), *Gramática descriptiva de la española* (vol. 3, cap. 59, 3805-3878). Espasa.
- Giorgi, A. y Longobardi, G. (1991). *The Syntax of Noun Phrases*. Cambridge University Press.s
- Gómez Belart, N., Colella, V. M. y Azzaro, V. M. (2023). *Programa de Lenguaje Claro: la experiencia de la Secretaría Legal y Técnica 2020-2023*. Gobierno de la Ciudad de Buenos Aires, Secretaría Legal y Técnica.
<https://gcba.github.io/programadelenguajeclaro/Programa.pdf>
- González Ruiz, R. (2008). Las nominalizaciones como estrategia de manipulación informativa en la noticia periodística: El caso de la anáfora conceptual. En I. Olza, M. Casado y R. González (Eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)* (pp. 247-259). Universidad de Navarra.
- Hartmann, P. (1955). *Nominale Ausdrucksformen im wissenschaftlichen Sanskrit*. Carl Winter Universitätsverlag.
- Hunt, K. (1965). *Grammatical structures written at three grade levels. National Council of Teachers of English Research Report N° 3*. National Council of Teachers of English.
- Hunt, K. (1970). Syntactic maturity in Schoolchildren and Adults. *Monographs of The Society for Research*, 35 (1), 1-67. doi: <https://doi.org/10.2307/1165818>
- Iturrioz, J. L. (1985). Abstracción sustantiva. Reificación de contenidos proposicionales. In J. L. Melena (Ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae* (vol. 1, pp. 395-414). Ediciones Universidad de Salamanca.
- Jacobi, H. (1903). Über den nominalen Stil des wissenschaftlichen Sanskrits. *Indogermanische Forschungen*, 14, 236-251.
- Lehmann, C. (1988). Towards a typology of clause linkage. In J. Haiman y S. A. Thompson (Eds.), *Clause Combining in Grammar and Discourse* (pp. 181-225). John Benjamins Publishing Co.
- Lehmann, C. (1995). *Thoughts on Grammaticalization*. Lincom Europa.
- López Samaniego, A. (2015). Etiquetas discursivas, hiperónimos y encapsuladores: Una propuesta de clasificación de las relaciones de cohesión referencial. *Rilce: Revista de Filología Hispánica*, 31(2), 435-462.
- López Samaniego, A. (2018). La encapsulación nominal en el discurso académico-científico oral y escrito: Patrones de aparición. *Caplletra Revista Internacional de Filología*, 64, 129-152. doi: <https://doi.org/10.7203/caplletra.64.11369>
- Nadal, L. (2024). El procesamiento de la encapsulación anafórica nominal y pronominal en hablantes de español L2: un estudio experimental con eyetracking. *Revista de Lenguas para Fines Específicos*, 30. doi: <http://dx.doi.org/10.20420/rufe/2024.697>
- NGLE = Real Academia Española y Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española*. Espasa.

- Picallo, M. C. (1999). La estructura del sintagma nominal: Las nominalizaciones y otros sustantivos con complementos argumentales. In I. Bosque y V. Demonte (Eds.), *Gramática descriptiva de la española* (vol. 1, cap. 6, 363-393). Espasa.
- Porzig, W. (1930-31). Die Leistungen der Abstrakta in der Sprache. *Blätter für Deutsche Sprachphilosophie*, 4, 66-77.
- Raible, W. (1992). *Junktion*. Carl Winter Universitätsverlag.
- Salvador Mata, F. (1985). Los índices de complejidad sintáctica, instrumentos de evaluación de la expresión escrita: estudio experimental en el ciclo medio de EGB. *Revista Interuniversitaria de Didáctica*, 3, 59-82.
- Sigaud-Sellos, P. (2010). *Aproximación a los conceptos de Legibilidad y Lecturabilidad: aplicación a la lectura de textos digitales*. Trabajo de investigación. Universidad de Navarra.



LA INCORPORACIÓN NOMINAL Y LA COMPOSICIÓN NOMINAL EN GRIEGO CLÁSICO: NOMOΘETHΣ COMO CASO DE ESTUDIO

Daniel AYORA¹
Universidad Autónoma de Madrid

Resumen

En los últimos años, se ha estudiado el impacto de la incorporación nominal en las lenguas. El griego antiguo es una lengua no polisintética en la que, en un verbo compuesto, se permite la incorporación del objeto directo del verbo base como primer elemento del compuesto, de modo que se forma un compuesto de rección verbal cuyo objeto directo transpone otro elemento del verbo simple de origen. Esta es la situación que la tipología de Mithun establece como un tipo II. Ahora bien, nos proponemos establecer en qué medida la incorporación nominal afecta a la composición nominal, en especial a los nombres de agente, y en qué medida esto ratifica las propuestas que se han realizado para los verbos compuestos en griego. *Palabras clave:* incorporación nominal; composición; nombres de agente; griego antiguo.

NOUN INCORPORATION AND NOMINAL COMPOUNDING IN CLASSICAL GREEK: NOMOΘETHΣ AS A CASE STUDY

Abstract

In recent years, the impact of Nominal Incorporation in languages has been studied. Ancient Greek is a non-polysynthetic language which, in complex verbs, allows the incorporation of the Direct Object of the base verb as the first element of the compound, thus forming a compound verb whose Direct Object transposes another element of the simple verb of origin. This is the situation that Mithun's typology establishes as Type II. Now, we propose to study Nominal Incorporation insofar as it affects nominal compounds, especially in agent nouns, and see to what extent it ratifies the proposals that have been given for Greek complex verbs.

Keywords: Noun Incorporation; Compounding; Agent Nouns; Ancient Greek.

RECIBIDO: 29/12/2024

APROBADO: 06/09/2025

1. danielayora93@gmail.com;  <https://orcid.org/0000-0003-4536-7191>

1. LA INCORPORACIÓN NOMINAL

En las últimas décadas, los lingüistas han estudiado y descrito la incorporación nominal (IN). Este tipo de construcciones han recibido diferentes explicaciones. Mithun (1986: 32) afirma: «in N[oun] I[norporation], as commonly understood since Sapir (1911), a noun stem is compounded with a verb stem to yield a more specific, derived verb stem». Por su parte, Baker (2002: 262) amplía esta definición para incluir en la incorporación nominal otro tipo de estructuras. En sus palabras: «a phenomenon in which two roots that would normally head distinct phrases are combined into a single morphological word». Mientras, para Murasugi (2014: 283) es «a construction in which a noun and a verb stem combine to yield a complex verb». De estas otras explicaciones, con matices, Luján (2025: 108) concluye que tienen en común que «in NI a nominal root combines with a verbal root and the resulting compound word is a verb that can be considered ‘derived’ from the original verbal root».

Este tipo de estructuras, como señala Baker (1998: 77-78), es frecuente en lenguas polisintéticas, como en yakui, una lengua uto-azteca. De esta recogemos el ejemplo de Haugen (2008: 118) para ilustrar el fenómeno:

- (1a) aapo maaso-ta peu-ta-k [sin IN]
 3SG ciervo-AC matar-INTR-PERF
 «Mató un ciervo.»
- (1b) aapo maaso-peu-te-n [con IN]
 3SG ciervo-matar-INTR-PAS
 «Estaba matando ciervos.»

En la oración de (1a) la forma verbal *peutən* tiene un objeto directo *maaso*, mientras que en la frase de (1b), *maaso* aparece como primer elemento de un verbo complejo cuyo segundo elemento es *peutən* al que se le han añadido formas específicas, como son los sufijos *-TAM*. Con todo, no son estructuras sinónimas. En (1a), la estructura tiene un valor perfectivo que no aparece en (1b). Es habitual en las lenguas incorporantes que haya diferencias entre la incorporación nominal y la estructura analítica. Estas diferencias suelen estar condicionadas por aspectos sintáctico-semánticos.

A este respecto, Mithun (1984) estableció una tipología de cuatro tipos para describir las funciones de la incorporación nominal². Esta tipología, según ella, está jerarquizada de manera que, si hay ejemplos del tipo II, también ha de haber casos del tipo I y así sucesivamente con los demás tipos³.

- a) Tipo I: similar a la composición léxica. El nombre incorporado (NI) está ligado estrechamente al verbo al que se incorpora y el compuesto hace referencia a una actividad frecuente, habitual o típica.
- b) Tipo II: con manipulación del caso. La incorporación nominal permite la promoción de un argumento oblicuo a la posición del caso dejada vacante por el nombre incorporado. Según Mithun (1984), se trata de la «extensión natural» del tipo I.
- c) Tipo III: con manipulación de la estructura del discurso. La incorporación nominal se emplea para los antecedentes conocidos del discurso anterior.
- d) Tipo IV: es «clasificatorio». Se incorpora al verbo una raíz sustantival general para acotar su ámbito; el verbo complejo puede ir acompañado de una frase sustantival que identifique el argumento al que se refiere el nombre incorporado.

Ahora bien, el griego antiguo es una lengua no incorporante, clasificación que suele darse a lenguas en las que se dan los tipos III y IV de Mithun. En este sentido, el griego antiguo permite una mínima incorporación, pero es preferible no darle dicha etiqueta. Trabajos como los de Luján (2025) han aplicado específicamente la tipología de Mithun (1984) sobre esta lengua y este autor ha llegado a la conclusión de que, claramente, el griego es una lengua de tipo II, ya que permite que el elemento nominal se configure como primer elemento del compuesto ($[N+V]_V$) y esta incorporación, que ha dejado libre la casilla de objeto directo del verbo, sea ocupada por un elemento, que responde al objeto del verbo complejo resultante, es decir, ese elemento se actualiza como objeto directo del verbo derivado de la incorporación⁴.

2. Otra tipología puede verse en Borik & Gehrke (2015: 2-4).

3. Recientemente se han introducido los conceptos de «pseudo-incorporación nominal» o simplemente «pseudo-incorporación» para describir un fenómeno similar: son casos en los que la IN implica frases que muestran un vínculo más estrecho de lo habitual con el verbo incorporante (Borik & Gehrke 2015).

4. Otras aproximaciones a la incorporación nominal en griego antiguo se pueden hallar en Pompei (2006), Asraf (2021) y Ricci (2025).

El hecho de que un verbo pueda sufrir un procedimiento morfológico por el cual su objeto directo es incorporado como primer elemento del compuesto no es nuevo, ya que es, en gran medida, el tipo de compuestos de rección verbal que ya describió Panīni para el antiguo indio y que los lingüistas han usado tradicionalmente para explicar los compuestos del indoeuropeo y, por consiguiente, del griego antiguo⁵.

Pero la intención de este trabajo es determinar en qué medida la incorporación nominal afecta a la composición nominal y no únicamente al verbo derivado de la incorporación. Concretamente, nos fijaremos en un tipo específico de derivados: los nombres de agente. Para exemplificar este fenómeno pondremos un compuesto en concreto, *νομοθέτης*⁶. Al comparar las características sintáctico-semánticas de esta palabra con otros nombres de agente veremos que la situación que describimos para *νομοθέτης* responde a la situación del conjunto de nombres de agente, y no a una palabra en particular.

Para ello, expondremos algunos problemas morfológicos de este compuesto (§2), veremos la opacidad o no al discurso que tiene el primer elemento del compuesto (tipo III) (§3) y analizaremos la complementación adnominal que recibe el compuesto (tipo II), para lo cual estableceremos la complementación de la colocación *νόμον τίθημι* (§4). Tras esto, recogeremos las conclusiones fundamentales a las que se ha llegado (§5).

2. LA MORFOLOGÍA DE NOMΟΘΕΤΗΣ

Normalmente, se ha interpretado *νομοθέτης* como un compuesto de rección verbal de *νόμος* 'ley' y *τίθημι* 'poner' (Bailly, Montanari, *CGL*). La combinación de estas dos palabras no es casual, ya que constituyen una colocación verbo-nominal léxica: *νόμον τίθημι* 'disponer una ley'⁷. De acuerdo con esto, *νομοθέτης* significa 'legislador' («lawgiver» LSJ, *CGL*; «legislator» Montanari). Esta explicación nos sigue pareciendo la más verosímil y no hacerlo proceder de *νομοθετέω*.

5. Cf. entre otros Risch (1974: 181-182), Meier-Brügger (2000: 272-275; 2017: 707) y Tribulato (2015).

6. Tras el estudio de los principales sufijos agentivos, se llega a la conclusión de que *νομοθέτης* es uno de los compuestos de rección verbal más atestiguado en la literatura clásica (*ca. 500x*, búsquedas en el *TLG*, cf. Ayora 2023: 743). Además, se puede comparar bien con las estructuras *νόμον τίθημι* y *νομοθετέω*, por lo que es el compuesto más ilustrativo para exponer esta cuestión que, como veremos, también afecta a otros sustantivos.

7. Para el concepto de colocación aplicado al griego antiguo cf. fundamentalmente Jiménez López (2016) y Baños *et alii* (2022) con amplia bibliografía. Para la colocación en griego antiguo, cf. *DiCoGrA s.v. νόμον τίθημι*.

De acuerdo con Mithun (1984: 854), podemos identificar *νομοθέτης* como un compuesto morfológico, esto es, una formación que el hablante percibe como una palabra única, aunque esté formada por varios elementos. Los principales motivos para interpretarlo así son:

- i) La palabra resultante tiene un único acento, que se rige por las leyes de acentuación habituales (*νομοθέτης*);
- ii) Como en el resto de nombres, estos compuestos reciben la marca flexiva únicamente en uno de los formantes, en caso del griego antiguo, en el segundo elemento, que constituye el final de la palabra (e.g. nom.sg. *νομοθέτης*, voc.sg. *νομοθέτης*, ac.sg. *νομοθέτης*);
- iii) Los compuestos tienen típicamente una vocal de unión -ο- entre ambos formantes; en el caso de *νομοθέτης* esta vocal coincide con la -ο- de la base *νομο-*, pero esto sucede claramente en otros compuestos, como aquellos cuyo primer elemento es un nombre femenino (e.g. ἐστροφής de ἐστρόφη), o un nombre de tema en -ι- (e.g. φυσιογνώμων de φύσις) o uno de tema en -σ- (e.g. ψευδοδότης de ψευδής);
- iv) Estos compuestos se conciben como una base unitaria para nuevos derivados. Para los verbos complejos se ponía el foco en que estas nuevas bases pueden ser, a su vez, preverbiadas para formar nuevos verbos. Aplicada esta cuestión a los compuestos nominales, los compuestos formados mediante incorporación pueden ser la base para formar verbos en -έω (e.g. ὀνομαθέτης ‘que da un nombre’ → ὀνομαθετέω ‘dar un nombre’)⁸. También pueden ser la base para crear otros nombres, ya sea mediante prefijación (e.g. ἀ-μετροπότης ‘que bebe sin medida’ con prefijo privativo ἀ-), ya sea mediante sufijación (e.g. διχόγνωμος ‘ambiguo’ → διχογνωμοσύνη ‘ambigüedad’, ἀγωνοθέτης ‘el que dispone los juegos, organizador de los juegos, juez en los juegos’ → ἀγωνοθετικός ‘relativo a la organización de los juegos’ y ἀγωνοθεσία ‘cargo de ἀγωνοθέτης, presidencia de los juegos’).

8. Con todo, sincrónicamente, estos derivados a partir de compuestos de rección verbal ($[N+V]_N \rightarrow V-έω$) pueden ser interpretados como verbos con incorporación nominal directamente, sin el paso por un compuesto nominal previo: $[X]_N [Y]_V \rightarrow [[X + Y]_{N/A} + -έω]_V$, cf. Pompei & Grandi (2012).

3. TRANSPARENCIA/OPACIDAD AL DISCURSO

Como decíamos, el tipo III de Mithun (1984) implica que el primer elemento del verbo complejo es susceptible de ser recogido y referenciado mediante fóricos, que nos aportan una información discursiva.

Como vemos en los casos, por ejemplo, de (2a) y (2b), el elemento incorporado (*vōμo-*) no es posible recogerlo mediante un fórico. En ambos casos, son frases en las que dicho elemento se ha de explicitar de nuevo: en (2a) a escasas palabras de distancia aparece *vōμoθéτης* frente a la colocación *τὸν νόμον ἔθηκεν*, mientras que en (2b) se opta por el predicado sintáctico, *vōμoθetéω* ‘disponer leyes’, por lo que se repite el elemento *vōμo-* tanto en el nombre como en el verbo.

(2a) Is.2.13

Ὀ γὰρ νομοθέτης, ὃ ἄνδρες, διὰ τοῦτο τὸν νόμον
ART PTCL legislador INTRJ hombre PRP eso ART ley.AC.SG
ἔθηκεν οὔτως,
dar.IND.AOR.3SG ADV

«Pues el legislador, oh hombres, por esto dispuso la ley así:»

(2b) Pl.Lg.880e

οἵς δὴ τοὺς νόμους ἐξ ἀνάγκης ὁ νομοθέτης
REL.DAT.PL.M PTCL ART ley.AC.PL PRP fuerza ART legislador
ἂν νομοθετοῖ, βουλόμενος αὐτῶν μηδέποτε
PTCL legislar.OPT.PRS.3SG querer.PTCP.PRS.NOM.SG.M ese ADV.NEG
χρείαν γίγνεσθαι.
necesidad resultar

«por quienes a la fuerza el legislador tenga que legislar leyes, sin querer necesitar nunca usarlas.»

Como se aprecia en estos ejemplos, no es un condicionante el que haya más o menos palabras entre ambos constituyentes, ni que el nombre esté en singular o plural, ya que no puede ser referenciado. Con todo, hay algún caso esporádico que podría hacernos dudar de este hecho. Por ejemplo, en (3) aparece el demostrativo *τούτους* y, aunque el compuesto está más próximo, es más fácil comprender que está referenciando a *νόμους*, que aparece previamente. Por tanto, en realidad no hay una referencialidad con respecto al elemento nominal de *vōμoθéτης*.

(3) Pl.Prt.326d⁹

ώς δε καὶ ἡ πόλις νόμους ὑπογράψασα, ἀγαθῶν καὶ
 ADV PTCL ADV ART ciudad ley.AC.PL inscribir.PTCL.AOR bueno CNJ
 παλαιῶν νομοθετῶν εὑρήματα, κατὰ τούτους ἀναγκάζειν καὶ
 antiguo legislador hallazgo PRP ese.AC.PL obligar CNJ
 ἄρχειν καὶ ἄρχεσθαι,
 gobernar.INF.PRS.ACT CNJ gobernar.INF.PRS.PAS
 «También como la ciudad, tras haber inscrito leyes, hallazgos de buenos y antiguos
 legisladores, según estas obliga a mandar y ser mandado.»

Otro caso es cuando el constituyente nominal está elíptico contextualmente. El único caso susceptible de ser así analizado es el pasaje de (4), en el que, claramente, θείς tiene como objeto directo el relativo οὖς, cuyo antecedente es νόμων, por lo que el elemento νόμους sigue evocado en la mente del hablante. Preferimos esta explicación a pensar que el hablante toma el elemento elíptico del primer elemento del compuesto νομοθέτης, que concuerda con θείς.

(4) Pl.Lg.772c

τότε δε ἀκίνητα θεμένους, ἥδη χρῆσθαι μετὰ
 ADV PTCL ADV poner.PTCP.AOR.IND.AC.PL.M CNJ ser_útil.INF.MED PRP
 τῶν ἄλλων νόμων οὖς ἔταξε κατ'
 ART otro.GEN.PL.M ley.GEN.PL.M REL.AC.PL.M establecer.AOR.3SG PRP
 ἄρχας ὁ θεὶς αὐτοῖς νομοθέτης.
 inicio ART poner.PTCP.AOR.NOM.SG él.DAT.PL.M legislador.NOM.SG.M
 «entonces establézcanse como inmutables y sean utilizadas con las demás leyes que
 estableció desde un principio el legislador que las había dispuesto para ellos.»

3.1. Comparación con νευροσπάστης

Νευροσπάστης ‘titiritero’ («puppet-show man» LSJ, «puppeteer» Montanari, CGL «moved by strings») es, literalmente, ‘el que tira de los nervios hacia sí’, y es otro compuesto de rección verbal a partir de νεῦρον ‘nervio’ y σπάω ‘tirar hacia sí’ (Bailly, Montanari, CGL).

Este compuesto da un ejemplo muy interesante para ver que el elemento incorporado al compuesto es opaco a la referencialidad del discurso. En (5) se describe

9. Véase también Pl.Lg.691b.

la acción del *νευροσπάστης* mediante el verbo base (preverbiado), ἐπισπάω, pero el objeto directo que toma este verbo no es *νεῦρον*, que es el elemento incorporado, ni tampoco un demostrativo que refiera a este, sino *μήρινθον*, otro nombre, próximo a la idea que refiere *νεῦρον*. Por lo que es otro ejemplo más en el que se evidencia que el elemento incorporado en los compuestos nominales es opaco a la referencialidad.

- (5) Arist.*de Mu.*398^b17

οἱ νευροσπάσται μίαν μήρινθον ἐπισπασάμενοι
ART titiritero.NOM.PL.M uno hilo.AC.SG tirar_arriba.PTCP.AOR.MED.NOM.PL.M
«Los titiriteros, tras haber tirado de un hilo,»

4. COMPLEMENTACIÓN ADNOMINAL

Νομοθέτης, como otros nombres en griego antiguo, puede tener un complemento adnominal en dativo¹⁰. Este elemento dependiente expresa, con respecto a *νομοθέτης*, el beneficiario, en tanto que es la persona, o grupo, para quien se dispone la ley¹¹. Casos de esta complementación son frecuentes, recogemos a modo de ejemplo el caso de (6)¹².

- (6) Arist.*Pol.*1274^a32

ἐγένετο δὲ καὶ Φιλόλαος ὁ Κορίνθιος
resultó PTCL ADV Filolao.NOM.SG.M ART corintio.NOM.SG.M
νομοθέτης Θηβαίοις.
legislador.NOM.SG.M tebano.DAT.PL.M
«Resultó también Filolao, el corintio, legislador para los tebanos.»

Este complemento en dativo transpone el tercer argumento de la construcción verbo-nominal léxica *νόμον τίθημι*. Esta colocación se halla 72 veces en Platón, por citar datos de un autor concreto. De estos 72 testimonios, 53 aparecen como una predicción absoluta, sin otros constituyentes oracionales, en 16 aparece con dativo y en 3 con el sintagma preposicional ἐπί-DAT. Exponemos el detalle de los datos en la tabla 1.

10. Para el dativo adnominal, *cf.* Díaz de Cerio (2022: 306).

11. Para la definición del beneficiario, *cf.* Revuelta (2014) con debate y bibliografía.

12. Véanse otros casos de *νομοθέτης* con dativo en Pl.*Lg.*635b, 680a, 729b; Arist.*Pol.*1274^b2, 24.

	Absoluto	Dativo	ἐπί-DAT	TOTAL
νόμον τίθημι	15	7	1	23
νόμους τίθημι	38	9	2	49
TOTAL	53	16	3	72

Tabla 1. Complementación de νόμον τίθημι en Platón

Como se aprecia en la tabla, la colocación es ligeramente más frecuente en plural (νόμους, 49x) que en singular (νόμον, 23x). Y, claramente, tanto en singular como en plural, la colocación tiende a expresarse de manera absoluta, sin ningún otro constituyente nuclear. Recogemos de nuevo en (7a) el pasaje que citábamos en (2a). Ahora bien, en un cuarto de los testimonios (26,8%), la colocación se construye con un beneficiario, que generalmente se expresa mediante dativo, como en (7b), pero que también puede ser expresado mediante un sintagma preposicional (7c).

(7a) Is.2.13

Ο γὰρ νομοθέτης, ὁ ἄνδρες, διὰ τοῦτο τὸν νόμον
 ART PTCL legislador INTRJ hombre PRP eso ART ley.AC.SG
 ἔθηκεν οὕτως,
 poner.IND.AOR.3SG ADV

«Pues el legislador, oh hombres, por esto dispuso la ley así:»

(7b) Pl.Mi.320b¹³

ὅθεν δὴ καὶ τὸν νόμους τούτους ἔθηκε τοῖς
 ADV PTCL ADV ART ley.AC.PL.M ese.AC.PL.M poner.IND.AOR.3SG ART
 αὐτοῦ πολίταις,
 él.GEN.SG.M ciudadano.DAT.PL.M

«desde entonces también dispuso estas leyes para sus ciudadanos»

(7c) Pl.Grg.488d

οἱ δὴ καὶ τὸν νόμους τίθενται ἐπὶ τῷ
 REL.NOM.PL.M PTCL ADV ART ley.AC.PL.M poner.IND.MED.PRS.3SG PRP ART
 ἐνὶ,
 uno.DAT.SG.M

«los que dictan leyes para un individuo,»

Por consiguiente, se puede concluir que el compuesto resultante de una incorporación nominal permite transponer al sintagma nominal el constituyente más

13. Véase también el ejemplo citado en (4) *supra*.

frecuente como complemento, el dativo. Aunque es cierto que el uso de *νομοθέτης* en los textos es, generalmente, sin complemento adnominal, es decir, que, como la colocación, tiende a aparecer en construcciones absolutas.

Son infrecuentes los casos en los que *νομοθέτης* tiene un nombre dependiente en genitivo en vez de un dativo (8). En estos casos, se ha integrado el tercer argumento de la colocación como segundo del nuevo compuesto, es decir, se dan las características que Mithun clasificó como tipo II para los verbos complejos. Esta es la situación que se ha descrito generalmente para los verbos en griego antiguo¹⁴ y vemos que la derivación no restringe este fenómeno. El genitivo, entre sus usos, permite transponer múltiples constituyentes oracionales. En este caso, el compuesto nominal incluye un tipo de integración sintáctica que no llega a darse en el verbo complejo *νομοθετέω*, que rige dativo (*cf. supra* 2b).

(8a) Pl.Lg.626a

τὸν	Κρετῶν	νομοθέτην
ART	cretense.GEN.PL.M	legislador.AC.SG.M
«el legislador de los cretenses»		

(8b) Arist.EN 1102^a11

Παράδειγμα	δὲ	τούτων	ἔχομεν	τοὺς	Κρητῶν
ejemplo.AC.SG.N	PTCL	esto.GEN.PL.N	tener.IND.ACT.PRS.1PL	ART	cretense.GEN.PL.M
καὶ	Λακεδαιμονίων	νομοθέτας			
CNJ lacedemonio.GEN.PL.M legislador.AC.PL.M					
«tenemos como ejemplo de esto a los legisladores de los cretenses y los lacedemonios»					

4.1. Comparación con otros compuestos agentivos

Pero este tipo de integración en el compuesto nominal no es lo más frecuente. Entre los nombres de agente en griego clásico *νομοθέτης* es uno de los pocos compuestos con incorporación nominal que pueda tener un genitivo objetivo. Vemos ahora los problemas de otros nombres.

Los ejemplos más ilustrativos son los de *όλβοδότης* ‘dador de riqueza’ (← *όλβος* ‘riqueza’ y *δίδωμι* ‘dar’), *πανδερκέτης* ‘que ve por completo’ (← *πᾶς* ‘todo’

14. Luján (2025).

y δέρκομαι ‘ver claramente’) y στρατηλάτης ‘comandante, líder de un ejército’ (← στρατός ‘ejército’ y ἐλάω ‘conducir’)¹⁵.

Estos compuestos incorporan como primer elemento el objeto directo del verbo base y tienen un genitivo que depende de ellos. Ahora bien, este es un genitivo de referencia, ya que expresa el dominio sobre el que la acción se da¹⁶. Esto es claro en el caso de στρατηλάτης, que está prácticamente lexicalizado y el genitivo τῶν Ἀχαιῶν en (9) es más fácil comprenderlo como referencial. Lo cual también ocurre con ὀλβοδότης. En (10), este compuesto con incorporación nominal tiene un complemento dependiente en dativo, por lo que mantiene la complementación del verbo base (δίδωμι). Por tanto, no se produce la incorporación sintáctica de que este tercer argumento verbal se identifique con la nueva segunda casilla del compuesto resultante. Al contrario, ὀλβοδότης, en este mismo pasaje, tiene un nombre dependiente en genitivo, εὐδαιμονίας, que hace referencia a una cualidad del primer elemento del compuesto.

(9) E.IA 801

τῶν	Ἀχαιῶν	ἐνθάδ'	ό	στρατηλάτης
ART	aqueo.GEN.PL.M	ADV	ART	comandante.NOM.SG.M

«allí el comandante de los aqueos (i.e. el que conduce el ejército de los aqueos)»

(10) E.Ba.572-573

τὸν	εὐδαιμονίας	βροτοῖς /	ολβοδόταν
ART	felicidad.GEN.SG.F	mortal.DAT.PL.M	riqueza_dador.AC.SG.M
πατέρ(α)			

padre.AC.SG.M

«padre que da riquezas de felicidad a los mortales».

El tercer caso es algo más complicado. La etimología de πανδερκέτης es clara, el problema es cómo interpretar el primer elemento del compuesto. La traducción del LSJ implica que se trataría de un elemento adverbial («seen by all»), por lo que se trataría de un compuesto de tipo determinativo, y no de rección verbal, pero si lo entendemos como el objeto de δέρκομαι, se trataría de otro caso de incorporación nominal con genitivo. Esta segunda interpretación es válida, ya que hay ejemplos de δέρκομai como el de (11) en donde este verbo tiene por objeto directo ἄπαντα

15. *Stricto sensu*, compuestos como δεσπότης o στρατηλάτης son, en origen, compuestos de rección, pero, sincrónicamente, están lexicalizados y han perdido esta proximidad semántica con su base, pues δεσπότης es ‘señor de la casa’ (IE *d(o)ms- ‘casa’ y *poti- ‘señor’, cf. DELG, EDG s.v. δεσπότης).

16. Cf. Martínez Vázquez *et alii* (1999: 116), Riaño (2022: 266).

‘todo’, sinónimo al elemento nominal del compuesto. Esta es la interpretación de Montanari («all-seeing») y el *CGL* («one who sees all»).

(11) B.15.51

Ζεὺς	ύψιμέδων	δς	ἄπαντα	δέρκεται
Zeus	alto_over	REL.NOM.SG.M	todo.AC.PL.N	ver.IND.PRS.3SG
«Zeus, alto soberano, el que todo lo ve»				

Si aceptamos, como Montanari y el *CGL*, que *πανδερκέτης* es un compuesto con incorporación nominal seguimos encontrando que puede tener un genitivo de referencia, como los que hemos visto en (9) y (10). En (12), el genitivo *βροτῶν* explica una parte del compuesto, pues Zeus ‘lo ve todo sobre los mortales’.

(12) E. El.1177-1178

Ζεῦ	πανδερκέτα /	βροτῶν
Zeus.VOC.SG.M	todo_ver.VOC.SG.M	mortal.GEN.PL.M
«Zeus, que todo lo ve sobre los mortales»		

5. CONCLUSIONES

Desde un punto de vista tipológico, el griego antiguo se clasifica como una lengua no polisintética. Se suele asumir que es una lengua no incorporante, aunque presente ciertos rasgos relacionados con la incorporación. En trabajos anteriores, se ha puesto el foco en explicar el comportamiento de los verbos complejos derivados tras la incorporación nominal, pero nosotros hemos querido aplicar esta teoría a la composición nominal. La muestra del trabajo se ha centrado en un grupo homogéneo de nombres, los nombres de agente en griego clásico y, dados los datos, hemos ejemplificado las características de estos compuestos mediante un nombre específico: *voμοθέτης*. Cuando hemos tenido ejemplos de otros compuestos nominales que aportaban información la hemos incluido.

Es bien sabido que hay cierto grado de incorporación nominal en los compuestos nominales del griego antiguo. Tradicionalmente se han denominado compuestos de rección verbal (o *Rektionkomposita*). En estos compuestos, se une un verbo base y un nombre, que corresponde al segundo argumento del verbo y que se incorpora como primer elemento del compuesto. Esta es la forma más básica de incorporación y que Mithun (1984) clasificó como de tipo I.

La evolución natural de este tipo I, como ella misma la denomina, es que la casilla que ha quedado liberada, ya que el objeto directo se ha incorporado como parte del compuesto, sea ocupada por otro complemento verbal, por lo que hay, también, una incorporación sintáctica.

Trabajos como los de Pompei (2006) y Asraf (2021) han puesto el foco sobre las características sintácticas y semánticas del tipo I para los verbos complejos en griego antiguo, mientras que Luján (2025) concluye que el griego alcanza hasta el tipo II, pero que no llega al tipo III, en el que el elemento nominal del compuesto puede ser referenciado por marcadores discursivos.

En este trabajo, hemos expuesto cómo la tipología jerarquizada de Mithun es válida también para explicar la composición nominal en griego antiguo. Por los datos analizados, las conclusiones afectan a los nombres de agente, por lo que es necesaria una investigación mayor que determine si esto afecta por igual también al resto de compuestos en griego antiguo.

Tras el estudio de los textos llegamos a la conclusión de que los compuestos de rección verbal son opacos a la referencialidad del discurso. Así, es necesaria la explicitación y repetición de *vóμον* (-ouς) en distintos casos, a pesar de que en la misma oración aparezca *vομοθέτης* (o *vομοθετέω*, el verbo derivado de él). Pero esta situación no es única de *vομοθέτης*. Como decimos, se pueden poner otros ejemplos que ilustran el mismo fenómeno, como hemos visto con *νευροσπάστης*. Por tanto, los compuestos nominales se comportarían del mismo modo, a este respecto, que los verbos complejos.

La última cuestión es si en la composición nominal se da la incorporación sintáctica de constituyentes en genitivo (tipo II de Mithun). Tras la descripción del comportamiento sintáctico de la colocación *vóμον τίθημι*, hemos visto la implicación que tiene en la transposición de sus argumentos en el compuesto nominal *vομοθέτης*. Esta colocación tiene un tercer argumento en dativo que expresa el beneficiario, y es un constituyente que se integra en el sintagma nominal generalmente en dativo. En menor medida hay testimonios de un genitivo adnominal que sí evidenciaría que se ha dado la integración sintáctica, por lo que los compuestos nominales podrían identificarse con el tipo II.

Ahora bien, los ejemplos de *vομοθέτης* con genitivo objetivo son escasos. En la mayoría de las ocasiones en las que un compuesto de rección verbal presenta un genitivo adnominal, este es un genitivo de referencia. Es la situación que hemos expuesto con *όλβοδότης*, *πανδερκέτης* y *στρατηλάτης*.

Queda, por tanto, profundizar en la incorporación nominal aplicada a otros compuestos nominales para establecer el alcance último de la cuestión que aquí se ha planteado y que sirve para sentar las bases de futuros trabajos.

AGRADECIMIENTOS

Agradezco al Prof. E.R. Luján y los revisores anónimos la lectura y comentarios a una versión anterior del trabajo. Con todo, los errores en que se pueda haber incurrido son solo atribuibles al autor de este trabajo.

FINANCIACIÓN

Este trabajo se ha realizado como parte de los proyectos de investigación PID2019-106606GB-C31 y PID2021-125076NB-C42 del Ministerio de Ciencia e Innovación del Gobierno de España.

REFERENCIAS

- Asraf, N. (2021). The Mechanism of Noun Incorporation in Ancient Greek. *Glotta*, 97, 36–72.
- Ayora, D. (2023). *Papeles semánticos y formación de palabras: los sufijos de nombres de agente en griego antiguo* [Tesis doctoral, Universidad Complutense de Madrid].
- Bailly = Bailly, A., Sechan, L., & Chantraine, P. (2000⁴). *Le Grand Bailly. Dictionnaire Grec Français*. Hachette.
- Baker, M. C. (1998). *Incorporation. A Theory of Grammatical Function Changing*. Chicago University Press.
- Baker, M. C. (2002). Incorporation. In W. Frawley (Ed.), *International Encyclopedia of Linguistics* (pp. 262–264). OUP.
- Baños, J. M., Jiménez López, M. D., Jiménez Martínez, M. I., & Tur, C. (Eds.) (2022). *Collocations in Theoretical and Applied Linguistics: From Classical to Romance Languages*. Guillermo Escolar Editor.
- CGL = Diggle, J. (Ed.) (2021). *The Cambridge Greek Lexicon*. CUP.
- DELG = Chantraine, P. et alii (2009²). *Dictionnaire étymologique de la langue grecque (Histoire des mots)*. Klincksieck.
- Díaz de Cerio, M. (2022³). Sintaxis y semántica del dativo. In M. D. Jiménez López (Coord.), *Sintaxis del griego antiguo. Volumen I: Introducción. Sintaxis nominal. Preposiciones. Adverbios y partículas* (pp. 297–325). CSIC.
- DiCoGrA = Jiménez López, M. D., Baños, J. M. (coords.). *DiCoGrA: Diccionario de colocaciones del griego antiguo*. <https://dicogra.iatext.ulpgc.es> [31/07/2025]
- EDG = Beekes, R. S. P. (2010). *Etymological Dictionary of Greek*. Brill.
- Haugen, J. D. (2008). *Morphology at the Interfaces. Reduplication and Noun Incorporation in Uto-Aztecan*. John Benjamins.
- Jiménez López, M. D. (2016). On Support Verb Constructions in Ancient Greek. *Archivio Glottologico Italiano*, 101(2), 180–204.

- LSJ = Liddell, H. G., Scott, R., & Jones, H. S. (1979⁹). *A Greek-English Lexicon*. Clarendon Press.
- Luján, E. R. (2025). Advances in Ancient Greek Linguistics. En J. de la Villa *et alii* (Eds.), *Proceedings of the International Conference on Ancient Greek Linguistics*. (pp. 107-126). De Gruyter Mouton.
- Martínez Vázquez, R., Ruiz Yamuza, E., & Fernández Garrido, M. R. (1999). *Gramática funcional-cognitiva del griego antiguo I: sintaxis y semántica de la predicación*. Universidad de Sevilla.
- Meier-Brügger, M. (2000). *Indogermanische Sprachwissenschaft*. Walter de Gruyter.
- Meier-Brügger, M. (2017). The Lexicon of Greek. In J. Klein, B. Joseph & M. Fritz (Eds.), *Handbook of Comparative and Historical Indo-European Linguistics*, vol. 1 (pp. 695-710). De Gruyter Mouton.
- Mithun, M. (1984). The evolution of noun incorporation. *Language*, 60.4, 847-894.
- Montanari = Montanari, F. (2015). *The Brill Dictionary of Ancient Greek*. Brill.
- Murasugi, K. (2014). Noun incorporation, nonconfigurationality, and polysynthesis. In A. Carnie, Y. Sato & D. Siddiqi (Eds.), *The Routledge Handbook of Syntax* (pp. 283-303). Routledge.
- Pompei, A. (2006). Tracce di incorporazione in greco antico. In P. Cuzzolin & M. Napoli (Eds.), *Fonologia e tipologia lessicale nella storia della lingua greca. Atti del VI incontro internazionale di linguistica greca* (pp. 216-237). Franco Angoli.
- Pompei, A., & Grandi, N. (2012). Complex -έο verbs in Ancient Greek. A case study at the interface between derivation and compounding. *Morphology*, 22.3, 399-416.
- Revuelta, A. (2014). Beneficiary. In G. K. Giannakis (Ed.), *Encyclopedia of Ancient Greek Language and Linguistics*, vol. 1 (pp. 230-234). Brill.
- Riaño, D. (2022²). Sintaxis y semántica del genitivo. In M. D. Jiménez López (Coord.), *Sintaxis del griego antiguo. Volumen I: Introducción. Sintaxis nominal. Preposiciones. Adverbios y partículas* (pp. 251-295). CSIC.
- Ricci, E. (2025). *Incorporazione del nome in greco antico: il caso di ποιέω 'fare'* [Sapienza Università di Roma].
- Risch, E. (1974³). *Wortbildung der homerischen Sprache*. Walter de Gruyter.
- Tribulato, O. (2015). *Ancient Greek Verb-Initial Compounds: Their Diachronic Development within the Greek Compound System*. De Gruyter.



GENERIC NULL OBJECTS IN SPANISH

Carlos MARTÍNEZ-GARCÍA¹
Universidad Complutense de Madrid

Abstract

This paper explores the grammatical constraints operating on Spanish generic null objects (e.g. *Esta película sorprende Ø* ‘this film surprises’), as well as their syntactic representation. Regarding the grammatical constraints, it is claimed in this paper that generic null objects in Spanish do not obtain their interpretation from an antecedent (i.e., they are non-anaphoric), must occur in generic sentences, be affected by the event denoted by the verb and receive a human reading. Also, their occurrence is lexically constrained: they only occur with predicates that allow human objects. Regarding their syntactic properties, generic null objects are not only semantically interpreted, but also syntactically represented, since they are sensitive to syntactic operations such as control, binding and secondary predication. Specifically, generic null objects are syntactically a set of D and φ -features (i.e., they display D, gender and number features). Thus, they are DPs. Regarding their analysis, it is proposed in this paper that generic null objects behave as variables bound by an operator in Spanish. This explains the different interpretations they receive depending on the operator in the sentence.

Keywords: animacy; generic null objects; genericity; *pro*; Spanish; variables.

LOS OBJETOS TÁCITOS GENÉRICOS EN ESPAÑOL

Resumen

Este trabajo explora las restricciones gramaticales que operan sobre los objetos tácitos genéricos del español (p. ej., *Esta película sorprende Ø*), así como su representación sintáctica. En cuanto a las restricciones gramaticales, en este trabajo se afirma que los objetos tácitos genéricos del español no obtienen su interpretación de un antecedente (es decir, no son anafóricos), deben aparecer en oraciones genéricas, estar afectados por el evento denotado por el verbo y recibir una lectura humana. Además, su aparición está restringida léxicamente: solo aparecen con predicados que legitiman objetos humanos. En cuanto a sus propiedades sintácticas, los objetos tácitos genéricos no solo se interpretan semánticamente, sino que

1. carlma27@ucm.es;  <https://orcid.org/0000-0002-8216-4759>

también se representan sintácticamente, ya que son sensibles a operaciones sintácticas como el control, el ligamiento y la predicción secundaria. En concreto, los objetos tácitos genéricos son sintácticamente un conjunto de rasgos D y φ (es decir, presentan rasgos D, de género y de número). Por tanto, se trata de SSDD. En cuanto a su análisis, en este artículo se propone que los objetos tácitos genéricos del español se comportan como variables ligadas por un operador. Este análisis explica las diferentes interpretaciones que pueden recibir según el operador que aparezca en la oración.

Palabras clave: animacidad; objetos tácitos genéricos; genericidad; *pro*; español; variables.

RECIBIDO: 24/02/2025

ACEPTADO: 05/12/2025

1. INTRODUCTION

In Spanish it is common to find examples like (1), where the transitive verbs *ayudar* ‘help’ and *emborrachar* ‘get drunk’ lack explicit objects. Contrary to Spanish indefinite null objects (see Campos, 1986; Brucart, 1999; Martínez-García, 2023 for information), the null objects in (1) do not require an antecedent to be interpreted (i.e., they are non-anaphoric). In this paper we discuss the grammatical conditions that allow null objects like those in (1).

(1) *Non-anaphoric null objects in Spanish*

- a. Un libro ayuda \emptyset_i
a.MASC.SG book help.IND.PRS.3SG
a [PRO_i pasar la tarde].
to spend.INF the.FEM.SG afternoon
'A book helps one to spend the afternoon.'
- b. El buen vino emborracha \emptyset poco.
the.MASC.SG good.MASC.SG wine make-drunk.IND.PRS.3SG a-bit
'Good wine gets one drunk just a bit.'

Although unnoticed in previous literature, null objects like those in (1) are possible in three syntactic contexts. In the first context, which appears in (2), they occur with psychological predicates. In this case, the object is affected and undergoes a change of state (it is an Experiencer).

(2) *Psychological predicates*

- a. La música clásica relaja \emptyset .
the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG
'Classical music relaxes one.'

- b. Los sonidos de la naturaleza calman \emptyset .
 the.MASC.PL sound.PL of the.FEM.SG nature calm.IND.PRS.3PL
 'Sounds of nature soothe one.'
- c. Las vacunas sanan \emptyset .
 the.FEM.PL vaccine.PL heal.IND.PRS.3PL
 'Vaccines heal one.'

The second context appears in (3), where the sentences are causative constructions in which the predicate selects an argumental small clause (see Gumié Molina, 2005, 2007; Landau, 2010 on this issue). This context shows more syntactic complexity than the one in (2).

(3) *Causative verbs*

- a. El dinero no hace \emptyset feliz.
 the.MASC.SG money not make.IND.PRS.3SG happy.SG
 'Money does not make one happy.'
- b. Su inteligencia deja \emptyset atónito.
 his/her intelligence leave.IND.PRS.3SG astonished.MASC.SG
 '{His/her} intelligence leaves one astonished.'
- c. Esta música pone \emptyset contento.
 this.FEM.SG music make.IND.PRS.3SG happy.MASC.SG
 'This music makes one happy.'

In the last context, which appears in (4), the null object controls the null subject of the subordinate clause. These are transitive structures with prepositional complements (Lebeaux, 1984; Rizzi, 1986; Authier, 1989, 1992; see Wurmbrand, 2002 for criticism). In (4), the null object controls PRO, as in versions with explicit objects (e.g., *La ambición conduce a uno a cometer errores* 'ambition leads one to make mistakes').

(4) *Control verbs*

- a. Esto lleva \emptyset_i a [PRO_i pensar lo siguiente].
 this.NEUT lead.IND.PRS.3SG to think.INF the.NEUT following
 'This leads one to think the following.'
- b. La ambición conduce \emptyset_i
 the.FEM.SG ambition lead.IND.PRS.3SG
 a [PRO_i cometer errores].
 to commit.INF error.PL
 'Ambition leads one to make mistakes.'

- c. Un libro ayuda \emptyset_i
 a.MASC.SG book help.IND.PRS.3SG
 a [PRO_{*i*} pasar la tarde].
 to spend.INF the.FEM.SG afternoon
 'A book helps one to spend the afternoon.'

As far as we are concerned, no description of the data in Spanish has been proposed in the literature (see Bosque & Gutiérrez-Rexach, 2009, pp. 359–363; Fábregas, 2023, pp. 25–28 for intuitions), but there is research on Italian and French, among other languages (see Kim, 1999 on Korean; Dvořák, 2017 on Czech).² Thus, the central goal of this paper is to explore the grammatical constraints operating on non-anaphoric null objects in Spanish and to describe their properties. Additionally, as a secondary goal, a preliminary analysis will be proposed. The description and the tentative analysis will rely on introspective data. Although it is true that the paper would benefit from further validation via corpus data or experimental methods, due to the novelty of the topic in Spanish, we consider that the incorporation of corpus data could be done in future research. For now, introspective data are enough to identify the basic grammatical properties of non-anaphoric null objects in Spanish.

This paper is organized as follows. First, we describe the previous analyses for non-anaphoric null objects in Romance, specifically for Italian and French (ch. 2). Second, we investigate the constraints on non-anaphoric null objects in Spanish (ch. 3). Third, given that non-anaphoric null objects are interpreted, but not pronounced, we discuss their syntactic representation and their grammatical properties (ch. 4). Next, we propose a tentative analysis relying on variables (ch. 5). Finally, we conclude with some final remarks (ch. 6).

2. PREVIOUS ACCOUNTS ON ROMANCE

Previous accounts for Romance non-anaphoric null objects are presented in this chapter. Specifically, we present the analysis in Rizzi (1986), which relies on pro_{arb} (sec. 2.1), and the analysis in Authier (1989, 1992), which relies on variables (sec. 2.2).

2. Other types of null objects (specifically, definite and indefinite ones) have received more attention in the literature, specially in Portuguese (see Barbosa, 2019, 2024) and Spanish (Campos, 1986; Clements, 2006; Laca, 2013; Verdecchia, 2022; Martínez-García, 2023, 2025), but, as they do not share basic properties with the ones studied in this paper (for instance, the former are anaphoric, while the latter are non-anaphoric), we leave them aside and treat the ones in this paper as an independent type of null object.

2.1. Rizzi (1986)

The first work on Romance non-anaphoric null objects is that of Rizzi (1986), which focuses on data from Italian. According to Rizzi (1986), the null objects in (5) are interpreted as ‘people, in general’ (i.e., they receive a generic, human interpretation). Additionally, they are said to be affected by the event denoted by the predicate. That is why they cannot occur as (affected) Themes, as explained in Rizzi (1986) (e.g., **Gianni vede felici* ‘Gianni sees happy’; see Rizzi, 1986, p. 539).

(5) *Italian*

Questo	conduce	Ø	alla	seguente	conclusione.
this.MASC.SG	lead.IND.PRS.3SG		to-the.FEM.SG	following	conclusion
‘This leads one to the following conclusion.’					

(Rizzi, 1986, pp. 503–507)

In Rizzi (1986), it is postulated that non-anaphoric null objects display the features [+GENERIC, +HUMAN]. These features attempt to explain the fact that they are interpreted as generic and human. That is why the null objects in (5) have been labelled as *generic null objects* (henceforth, GNOs) in the literature (Cummins & Roberge, 2004; Dvořák, 2017). Also, in Rizzi (1986), non-anaphoric null objects have been labelled as *arbitrary null objects*, where the term *arbitrary* refers to arguments displaying human interpretations (see Jaeggli, 1982, 1986; Suñer, 1983; Hernanz, 1989; Kim, 1991; Masullo, 2015 for discussion). The arbitrary label corresponds to a human variant of genericity, so that «arbitrary interpretations are a general indefinite referring to persons» (Krifka et al., 1995, p. 124).

In Rizzi (1986), the question is raised as to which empty category the GNO in (5) is. The options in Government and Binding are *pro*, PRO, the variables and the traces (Chomsky, 1981). Since GNOs are non-anaphoric, PRO and the traces are ruled out (they are anaphoric by hypothesis). Thus, according to Rizzi (1986), GNOs display the feature [–ANAPHORIC]. Additionally, GNOs cannot behave as variables, since variables are [–PRONOMINAL]. According to Rizzi (1986), Italian GNOs are subject to pronominal constraints like Principle B, since any GNO allows to be referentially linked to a phrase that c-commands the GNO and that is outside the category governing the GNO. As shown in (6), the first PRO is outside the category governing the GNO.

(6) *Italian*

È difficile [PRO' sperare [che il
 be.IND.PRS.3SG difficult expect.INF that the.MASC.SG
 governo possa autorizzare Ø
 government AUX.SBJV.PRS.3SG authorize.INF
 a [PRO" vivere così]].
 to live.INF like-this

'It is difficult to expect that the government can authorize one living like this.'

(Rizzi, 1986, p. 512)

This type of propositions can receive two readings. Specifically, (6) can mean 'it is difficult for x to hope that the government can authorize x to live like this' (i.e. $\text{PRO}' = \text{PRO}''$) or 'it is difficult for x to hope that the government can authorize y to live like this' (i.e. $\text{PRO}' \neq \text{PRO}''$). That is, GNOs can be ambiguous between two interpretations because, according to Rizzi (1986), they behave as *pro*. This implies that, in a context where a phrase falls within the category governing the GNO, the example should not allow the reading of the phrase governing the GNO and the reading of GNO to coincide referentially. In (7), where the first PRO is within the category governing GNO, only the reading 'in this department, it is difficult for x to force y to work' (i.e., $\text{arb}' \neq \text{arb}''$) is available. This is the behavior expected of a pronominal element, according to Rizzi (1986).

(7) *Italian*

In questo dipartimento, [PRO_{arb'} costringere Ø_{arb''}]
 in this.MASC.SG departament force.INF
 a [PRO" lavorare] è difficile.
 to work.INF be.IND.PRS.3SG difficult

'In this department, forcing one to work is difficult.'

(Rizzi, 1986, p. 512)

In conclusion, according to Rizzi (1986), the features of GNOs in Italian are [+PRONOMINAL, -ANAPHORIC], which are those of *pro*. Additionally, GNOs display the features [+GENERIC, +HUMAN], contrary to classic *pro*. This difference is captured by the feature *arbitrary* ($\text{pro}_{\text{arb}'}$) in Rizzi (1986). In essence, the label *arbitrary* adds the features [+GENERIC, +HUMAN] to the features that characterize *pro*. Rizzi's (1986) analysis appears in (8).

(8) Rizzi's (1986) analysis

$$[\text{VP} [\text{V} \text{ costringere}] [\text{DP} \text{ } \text{pro}_{\text{arb}}]]$$

In the following section we explore the analysis in Authier (1989, 1992), which, contrary to that in Rizzi (1986), relies on variables.

2.2. Authier (1989, 1992)

Contrary to Rizzi (1986), Authier (1989, 1992) points out that GNOs in French, as those in (9), do not display pronominal behavior. To illustrate that, two tests are presented in Authier (1989, 1992): the quantificational force of GNOs and the so-called 'equative structures' (i.e., sentences where two entities are equated with each other; see Lebeaux, 1984; Authier, 1992 for an overview).

(9) *French*

L' ambition amène Ø à commettre des erreurs.
 the.FEM.SG ambition lead.IND.PRS.3SG to commit.IND.PART error.PL
 'Ambition leads one to commit errors.'

(Authier, 1989, pp. 46–47)

According to Authier (1989, 1992), in languages like Spanish, the interpretation of null subjects can be ambiguous between a definite and an arbitrary reading. As shown in (10), *Llaman a la puerta* can mean 'they knock at the door' (definite reading), where the null subject is interpreted as a definite pronominal, as in (10a), and 'someone knocks at the door' (arbitrary reading), as in (10b), where the null subject is interpreted as an arbitrary pronominal (see Jaeggli, 1986; Hernanz, 1988 on this point). Thus, the arbitrary reading is related to some kind of 'existential' quantification.

(10) *Definite and arbitrary readings of null subjects*

a. *pro* Llaman a la puerta.
 knock.IND.PRS.3PL to the.FEM.SG door

'They knock at the door.'

b. *pro_{arb}* Llaman a la puerta.
 knock.IND.PRS.3PL to the.FEM.SG door

'Someone knocks at the door.'

According to Authier (1989, 1992), while the proposition in (9), instead of receiving the existential reading in (11a), receives the reading in (11b), the reading of the proposition in (10b) is close to an existential one, like the one in (11a) (see Dobrovie-Sorin, 1997 on this issue). In other words, GNOs receive generic readings, while arbitrary null subjects only receive an existential reading. This asymmetry regarding the quantificational force between GNOs and arbitrary null subjects in Spanish leads Authier (1989, 1992) to conclude that the empty categories in (9) and (10) are not the same (i.e., GNOs cannot be analyzed as an arbitrary pronominal element). Hence, French GNOs do not display pronominal behavior (pace Rizzi, 1986).

(11) *Existential and generic readings*

a. *Existential*

'There is some x , x is a person, such that the ambition leads x to x commit errors.'

b. *Generic*

'For any x , x is a person, such that the ambition leads x to x commit errors.'

The second argument in Authier (1989, 1992) relies on equative structures (see Lebeaux, 1984 on this point). Authier (1989) follows Jaeggli (1982, 1986) explaining that, when two arbitrary subjects converge in Spanish, both can have a different reference, as in (12), where the people who earn money are different from those who control drug trafficking.

(12) *Arbitrary readings of null subjects*

Para	que <i>pro</i> _{arb}	puedan	ganar	tan	dinero
for	that	AUX.SBJ.PRS.3PL	earn.IND	so-much	money
vendiendo	drogas	debe	ser	que <i>pro</i> _{arb}	
sell.GER	drug.PL	AUX.IND.PRS.3SG	be.IND	that	
controlan	muy	poco	estrictamente	el	tráfico
control.IND	very	little	strictly	the.MASC.SG	trade
de	drogas	en	este		
of	drug.PL	in	this.MASC.SG	páis.	

'For them to be able to make so much money selling drugs, they must have very little control over the drug trade in this country.'

(Authier, 1989, p. 49)

This leads Authier (1989) to conclude that, if the empty category in (13) is the same as that of French GNOs, GNOs should be able to refer to different entities, such as

Spanish arbitrary subjects (which display pronominal behavior). However, if a sentence displays two GNOs in French, the reference of the GNOs must be the same. This is clear when considering the equative structure in (13), which is contradictory: the same person cannot be intelligent and stupid at the same time. Consequently, Authier (1989, 1992) proposes that French GNOs should not be analyzed as *pro* (pace Rizzi, 1986).³

(13) *French*

#Une	thérapeutique	qui	rend Ø	intelligent
a.FEM.SG	therapy	that	make.IND.PRS.3SG	intelligent.MASC.SG
est	une	thérapeutique		
be.IND.PRS.3SG	a.FEM.SG	therapy		
qui	rend Ø	stupide.		
that	make.IND.PRS.3SG	stupid.MASC.SG		

‘A therapy that makes one smart is a therapy that makes one stupid.’

(Authier, 1989, pp. 49–50)

Having shown that French GNOs do not display pronominal behavior, Authier (1989, 1992) proposes that GNOs are analyzed as variables. The first argument has to do with the idea that French GNOs are subject to weak crossover effects. In (14a), the possessive pronoun and the GNO share the arbitrary reference. However, if the pronoun appears to the left of the sentence, as in (14b), both elements cannot share reference (the possessive can only be co-referential with an ostensibly accessible referent). This evidence suggests that French GNOs violate weak crossover effects, implying that French GNOs are variables, akin to traces bound by an operator, which can only bind what lies to its ‘right’ in the structure.

(14) *French*

a.	La	chase	rend Ø _{arb}	
	the.FEM.SG	hunting	make.IND.PRS.3SG	

3. Interestingly, as correctly pointed out by an anonymous reviewer, modality could play an important role in sentences with GNOs. Specifically, while the sentence *#Una terapia que pone triste es una terapia que pone alegre* ‘A therapy that makes one sad is a therapy that makes one happy’ is odd, a sentence with a different modality, such as *Una terapia que pone triste puede ser una terapia que pone alegre* ‘A therapy that makes one sad could be a therapy that makes one happy’, seems somewhat less odd. The asymmetry between these sentences has to do with the fact that, while the former is a true equative structure, the latter is not, as it introduces the modality verb *poder* ‘can’. Hence, the former introduces an assertion about a concrete state of affairs in the world (thus entailing the contradiction), while the latter only introduces an assertion about the possibility of a concrete state of affairs in a possible world (thus not entailing a real contradiction).

amoureux de son_{arb/i} chien.
in-love of one's dog

'Hunting leaves one in love with one's dog.'

b. Son_{?arb/i} chien rend Ø_{arb} amoureux de la chasse.
one's dog make.IND.PRS.3SG in-love of the.FEM.SG hunting
'His dog leaves one in love with hunting.'

(Authier, 1989, p. 50)

The second argument has to do with scope ambiguities. It has been considered that in sentences with two or more quantified phrases, ambiguities of scope can arise, depending on whether one quantifier is 'above' or 'below' the other one. Since May (1977), it is known that this phenomenon arises as an effect of quantifier raising, which operates in LF.

Given that quantified sentences leave variables where they were generated, this leads Authier (1989) to think that, if GNOs are variables, their occurrence in quantified sentences should entail scope ambiguities. Authier (1989) explains that two interpretations are possible in (15): it can be interpreted as 'for some x , x is a thing, and for all y , y is a person, x leads y to break the rules if y is cunning' (i.e., the existential quantifier is 'above' the universal one) and as 'for all y , if y is a person, there exists some x , such that x is a thing and x leads y to break the rules if y is cunning' (i.e., the universal quantifier is 'above' the existential one).

(15) *French*

Dans ce camp militaire, quelque chose
in this.MASC.SG camp military any thing
pousse Ø à [PRO enfreindre le règlement]
make.IND.PRS.3SG to break.INF the.MASC.SG rules
quand on est faux-jeton.
when one be.IND.PRS.3SG cunning

'In this military camp, something induces one to break the rules when one is cunning.'

(Authier, 1989, p. 53)

According to Authier (1989), this is a compelling reason to conclude that GNOs behave as variables, since variables are bound by an operator. The different ways in which the variable can be bound by the operator gives rise to the two interpretations in (16).

(16) *Scope ambiguity*

- a. $\exists x, x$ is a thing, such as x induces $\forall y, y$ is a person, to break the rules if y is cunning.
- b. $\forall y, y$ is a person, $\exists x, x$ is a thing, such as x induces y to break the rules if y is cunning.

(Authier, 1989, p. 53)

The quantificational force of GNOs depends on the quantificational force of the quantifier in the sentence. According to Authier (1989), (17a) is interpreted as ‘normally, fear makes many people to run away’, while (17b) reads as ‘this medicine never makes anyone sick’.

(17) *French*

- a. Souvent, la peur pousse \emptyset à fuir.
normally the.FEM.SG fear make.IND.PRS.3SG to run-away.INF
'Normally, fear forces one to run away.'
- b. Ce médicament ne rend jamais malade.
this.MASC.SG medicine not make.IND.PRS.3SG never sick
'This medicine never makes one sick.'

(Authier, 1989, p. 55)

Specifically, the interpretation of the GNOs in (17) depends on the type of quantifier that has scope over them. Since the quantifier in (17a) is *souvent* ‘normally’, the GNO is interpreted as ‘many people’, and, since the quantifier of (17b) is *jamais* ‘never’, the GNO is interpreted as ‘anyone’. Thus, Authier’s (1989) analysis relies on representing the GNOs as free variables, as in (18), where e represents the variable. Thus, GNOs are generated as free variables bound in LF by an operator. This operator can be explicit, as *souvent* ‘always’ in (18a), or implicit, as in (18b) (see Lewis, 1975 on this point).

(18) *Authier’s (1989) analysis*

- a. $\text{Souvent}_i [\text{VP} [\text{V} \text{ pousse}] \llbracket_{\text{DP}} e_i]]$
- b. $\text{OP}_i [\text{VP} [\text{V} \text{ pousse}] \llbracket_{\text{DP}} e_i]]$

In the following chapter we present the grammatical constraints on GNOs in Spanish so that we can propose which analysis better explains the facts.

3. GRAMMATICAL CONSTRAINTS

In this chapter we present the grammatical conditions that allow GNOs in Spanish. They are related to genericity (sec. 3.1), animacy (sec. 3.2), and affectedness (sec. 3.3).

3.1. *Genericity*

It has been pointed out in the literature that the licensing of GNOs is related to genericity (see Rizzi, 1986 on Italian; Authier, 1989, 1992 on French; Dvořák, 2017 on Czech). In this section we test whether this holds for Spanish GNOs.

Genericity is understood as a semantic property that characterizes sentences that do not refer to concrete instances of states of affairs in precise spatio-temporal coordinates. A *generic sentence* is a sentence in which «the property described by the verbal predicate is an ‘essential’ property of some entity mentioned in the sentence» (Krifka et al., 1995, p. 19). In (19a), the property of the subject is interpreted as ‘render deaf’, that is, ‘for every *x* who hears this noise *y*, every *x* will become deaf’. The same can be said of (19b), where definite articles appear, given that they enhance the genericity of the sentence (Leonetti, 1990, 1999).

(19) *Genericity and GNOs*

- a. Este ruido deja Ø sordo.
 this.MASC.SG noise make.IND.PRS.3SG deaf.MASC.SG
 ‘This noise deafens one.’
- b. En aquél país los monarcas
 in that.MASC.SG country the.MASC.PL king.PL
 castigan Ø con dureza.
 punish.IND.PRS.3PL with harshness
 ‘In that country, the kings punish one severely.’

A similar reading is obtained with the second person clitic *te* ‘(to) you’, which in Spanish is the person that can also mark genericity, as in (20) (Hernanz, 1988).

(20) *Genericity and clitics*

- a. Este ruido te deja sordo.
 this.MASC.SG noise to-you make.IND.PRS.3SG deaf.MASC.SG
 ‘This noise deafens you.’
- b. En aquél país los monarcas
 in that.MASC.SG country the.MASC.PL king.PL

te castigan con dureza.
 to-you punish.IND.PRS.3PL with harshness
 'In that country, the monarchs punish you harshly.'

One of the factors favoring sentential genericity is that the inflection of the matrix verb is imperfective (Krifka et al., 1995). If the sentences in (20) appear in perfective tense, they are unnatural, as in (21), where the verbs are inflected in perfective tenses.

(21) *Grammatical aspect and GNOs*

- a. *Este ruido dejó Ø sordo.
 this.MASC.SG noise make.IND.PRF.3SG deaf.MASC.SG
- b. ??En aquel país los monarcas
 in that.MASC.SG country the.MASC.PL king.PL
 han castigado Ø con dureza.
 AUX.IND.PRF.3PL punish.IND.PRS.3PL with harshness

The ungrammaticality of (21) is due to the fact that GNOs must occur in generic sentences, typically sentences with imperfective verbs. This also favors the habitual reading of propositions with GNOs. The habitual reading consists of linking a property to all situations in which the event designated by the proposition occurs. (21b) links the property 'punish severely' to a situation, which is 'in this country'. In Krifka et al. (1995), habituals consist of generic operators that bind variables. The reading in (20b) is represented in (22).

(22) *Semantics of generics*

- ... Gen_{*s,x*} [los monarcas castigan a *x* en *s*] [en aquel país *x* en *s*]

Genericity is not encoded in a single element, but is given by the confluence of different 'ingredients'. GNOs are one of such ingredients, but so are determiners, subject characterizing readings, and imperfective tenses (see Krifka et al., 1995 for discussion). Thus, what is generically interpreted in (20a) is the GNO, whereas in (20b) it is the GNO and the sentence. This suggests that GNOs are possible as long as the object is interpreted generically. In fact, it is possible to find cases where the sentence is not generic, but the null object is (e.g., *Esta película sorprendió Ø en su época* 'this film surprised at its time', where, despite having a verb in perfective tense, it is denoted that the action of surprising took place several times in the

space-time denoted by *en su época* ‘at its time’). Spanish shares these restrictions with Italian (Rizzi, 1986) and French (Authier, 1989, 1992).

As is explained in Chierchia (1998) and Krifka et al. (1995), genericity can be linked to a sentence or to a phrase. (23) means ‘the noise of the accordion deafens people, in general’ (i.e., the object is interpreted generically). Likewise, the subject is also interpreted generically, in the sense that any noise emitted by any accordion leaves people deaf. Furthermore, *el ruido del acordeón* ‘accordion noise’ does not refer to a particular entity, but to all entities that can be understood as ‘accordion noise’. That is, the generic reading is obtained by the combination of subject genericity and object genericity. However, in (20a), *este ruido* ‘this noise’ refers to a particular noise.

(23) *Subject and object genericity*

- El ruido del acordeón deja Ø sordo.
 the.MASC.SG noise of-the.MASC.SG accordion make.IND.PRS.3SG deaf.MASC.SG
 ‘Accordion noise deafens one.’

This generic reading is not possible with a third person clitic. If this clitic occurs, a reading that particularizes the event denoted by the predicate in an individual appears (e.g., *El ruido del acordeón lo deja sordo* ‘accordion noise deafens him’). The alternation between a GNO and a third person clitic is not possible (two interpretations are obtained). Thus, the generic reading is not obtained through sentential genericity nor through subject genericity: it is also necessary that the predicate selects a null object that favors this reading.

So far, it has been explained that GNOs in Spanish receive generic readings. That is, sentences with GNO seem to be universally quantified. The data in (24), with explicit objects, denote events particularized to an individual or set of individuals, whereas the events designated by the predicates in (25) are applicable to any individual.

(24) *Object particularizing readings*

- a. La música clásica relaja a Juan.
 the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG DOM Juan
 ‘Classical music relaxes Juan.’
- b. Las vacunas sanan
 the.FEM.PL vaccine.PL heal.IND.PRS.3PL
 a la mujer de mi primo.
 DOM the.FEM.SG wife of my cousin.MASC.SG
 ‘Vaccines heal my cousin’s wife.’

(25) *Object generic readings*

- a. La música clásica relaja Ø.
 the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG
 'Classical music relaxes one.'
- b. Las vacunas sanan Ø.
 the.FEM.PL vaccine.PL heal.IND.PRS.3PL
 'Vaccines heal one.'

In (24a) it is interpreted that music relaxes only Juan, while in (25a) it follows that music relaxes everyone who listens to it. The same can be applied to the other contrasts. Thus, it seems that the propositions with GNOs are universally quantified (Authier, 1989, 1992). The interpretations of each of them correspond to the paraphrases of (26).

(26) *Universal readings*

- a. For any x , if x is a person, classical music relaxes x .
 b. For any x , if x is a person, vaccinations heal x .

However, there are reasons to believe that universal quantification is not involved in generic sentences with GNOs. If the sentences in (25) were universally quantified, they would behave similar to the versions with the universal quantifier *todo* 'all' in (27). Specifically, if universal quantification were involved, predicates with GNOs should not allow exceptions (Krifka et al., 1995). However, propositions with GNOs can allow exceptions, whereas universally quantified propositions do not allow exceptions, according to Krifka et al. (1995): if classical music does not relax only one person in the world, the proposition in (27a) is still true, but the proposition in (25a) is not.

(27) *Universal quantification*

- a. La música clásica relaja
 he.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG
 a todo el mundo.
 DOM all.MASC.SG the.MASC.SG world
 'Classical music relaxes everyone.'
- b. Las vacunas sanan
 the.FEM.PL vaccine.PL heal.IND.PRS.3PL
 a todo el mundo.
 DOM all.MASC.SG the.MASC.SG world
 'Vaccines heal everyone.'

However, this difference does not seem to be definitive, since the propositions in (27) admit exceptional clauses, as in (28a), while the propositions in (25) do not, as in (28b). This is due to the fact that, while the speaker is not part of the set of individuals designated by universal quantification in (27), it is part of the set of individuals designated by the GNO in (25). That is, the judgment in (27) applies to all humans, with the possibility of excepting the speaker, but the judgment in (25) includes all humans, including the speaker. This prevents the speaker from being 'excepted'.⁴

(28) *Exceptional clauses*

- a. La música clásica relaja
the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG
a todo el mundo, excepto a mí.
DOM all.MASC.SG the.MASC.SG world except DOM me
'Classical music relaxes everyone, except me.'
- b. *La música clásica relaja Ø, excepto a mí.
the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG except DOM me

Another difference has to do with the ability of (25) and (26) to allow restrictions on the domain of predication. These restrictions consist in reducing the set of entities to which the generalization denoted by the proposition applies. If the domain of predication is restricted, GNO structures do not express adequate generalizations with respect to that domain restriction, but universally quantified sentences do, as pointed out in Krifka et al. (1995). In a situation where, in the results of a vaccine experiment, tests of one hundred vaccines have been found to be 100% positive for healing the people on whom the doses have been tested, only one hundred vaccines are involved, not all the existing vaccines in the world. That is, there is a restriction of the domain of application of the generalization in (25b) and (27b). In this interpretation, the proposition in (25b) is true, but the proposition in (27b) is not. The same can be applied to the other contrasts. It follows that universally quantified propositions can fit domain reductions, whereas propositions with GNOs cannot. Thus, genericity should not be confused with universal quantification in these cases.

4. Interestingly, as correctly pointed out by an anonymous reviewer, it is not only the case that the speaker cannot be excluded in sentences with GNOs in Spanish; other participants cannot not be excluded either (e.g., **La música clásica relaja Ø, excepto a mis hijos* 'Classical music relaxes everyone, except me'). This reveals a clear syntactic asymmetry between universal quantified sentences and sentences with GNOs: while the former allow for exceptional clauses, the latter do not.

3.2. Animacy

Another property of GNOs is their human interpretation, which they share with generic subjects (Jaeggli, 1982, 1986; Hernanz, 1989; Masullo, 2015). The GNOs in (29) are interpreted as human generic entities. As said above (see ch. 2), the same interpretation has been described for GNOs in French (Authier, 1989, 1992) and Italian (Rizzi, 1986).

(29) *Animacy and GNOs*

- a. La música clásica relaja Ø.
the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG
'Classical music relaxes one.'
- b. Las vacunas sanan Ø.
the.FEM.PL vaccine.PL heal.IND.PRS.3PL
'Vaccines heal one.'

The human interpretation of GNOs is evidenced by control (Lebeaux, 1984; Roberge, 1991). In (30a), the GNO controls the interpretation of the null subject (PRO) of the non-finite sentence *abandonar la ciudad* 'leave the city'. If the object denotes a non-human entity, as in (30b), PRO is interpreted as non-human too. This restriction explains why they cannot occur with predicates that do not select human objects (e.g., **Aquí fabrican Ø bien* 'here they manufacture well').

(30) *Control verbs and GNOs*

- a. La lluvia obligaba Ø
the.FEM.SG rain force.IND.IPFV.3SG
a [PRO abandonar la ciudad].
to abandon.INF the.FEM.SG city
'The rain forced people to leave the city.'
- b. La lluvia hizo a los perros
the.FEM.SG rain make.IND.PRS.3SG to the.MASC.PL dog.MASC.PL
[PRO abandonar la ciudad].
abandon.INF the.FEM.SG city
'The rain made the dogs leave the city.'

Due to the human interpretation of the GNO in (30a), the null subject (PRO) of the sentence *abandonar la ciudad* 'leave the city' must be interpreted as human. That is, the null subject of the subordinate sentence gets its interpretation from the GNO.

As animacy is involved, another factor for licensing GNOs is related to the selectional properties of the verb. Given that GNOs receive a human interpretation, verbs such as *preocupar* ‘worry’, *insultar* ‘insult’, *premiar* ‘reward’, or *castigar* ‘punish’ have in common the need for the object to be human (Martínez-García & Romero Heredero 2024). It is common to find verbs that (typically) select human objects. This would explain the contrast in (31). While *pegar* ‘hit’ selects only human objects (e.g., *Pegar a un compañero* ‘hit a partner’, but **Pegar (a) una mesa* ‘hit a table’), *golpear* ‘hit’ can select animate (e.g., *Golpear a un amigo* ‘hit a friend’; *golpear (a) una vaca* ‘hit a cow’) and inanimate (e.g., *Golpear la mesa* ‘hit the table’) objects. Hence, GNOs are allowed in (31a), but not in (31b).

(31) *Lexical constraints on GNOs*

- a. Siempre pegaban Ø a la salida de ese colegio.
always hit.IND.IPFV.3PL at the.FEM.SG exit of that.MASC.SG school
'They always hit one at the exit of that school.'
- b. ??Siempre golpeaban Ø a la salida de ese colegio.
always hit.IND.IPFV.3PL at the.FEM.SG exit of that.MASC.SG school
'They always hit one at the exit of that school.'

(Martínez-García & Romero Heredero, 2024, p. 271)

The selectional properties of the predicate ‘make it easier’ to obtain the human interpretation of the object in the first case than in the second. Thus, the GNO is acceptable in (31a), but not in (31b), since the predicate in (31b) does not select only human objects, unlike the predicate in (31a).

3.3. Affectedness

Finally, GNOs in Spanish are affected by the event denoted by the predicate. Intuitively, a participant of an event is considered to be affected when it suffers or experiences some change of state (physical or mental) as a result of the event (see Beavers, 2008 for discussion). The object is affected when it receives a thematic role other than (affected) Theme (see Rizzi, 1986 on this issue). Examples (32a–b), where GNOs present thematic roles as Experiencer, Beneficiary, and Goal (Dowty, 1991), are possible, Themes such as the ones in (32c–d) are excluded.⁵

5. As correctly pointed out by an anonymous reviewer, a potential contrast that would support the claim that Spanish GNOs are human and affected could be *A veces el dolor rompe* ‘Sometimes pain

(32) *Affectedness and GNOs*

- a. A veces Mario {impresiona / preocupa} Ø.
sometimes Mario impress.IND.PRS.3SG worry.IND.PRS.3SG
'Sometimes Mario {impresses / worries} people.'
- b. A veces el director {castiga / premia}
sometimes the.MASC.SG director punish.IND.PRS.3SG reward.IND.PRS.3SG
Ø sin razón aparente.
without reason apparent
'Sometimes the director {punishes / rewards} people for no apparent reason.'
- c. *A veces Mario {conoce / sabe / contempla} Ø.
sometimes Mario know.IND.PRS.3SG contemplate.IND.PRS.3SG
- d. *Mi hermano {admira / aprecia} Ø.
my brother.MASC.SG admire.IND.PRS.3SG appreciate.IND.PRS.3SG

Consequently, the verb that allows for a GNO must impose on it a change, whereby the object becomes affected by the action. The assignment of a thematic role different from Theme, together with the affected character of the object, explains the ungrammaticality of (32c-d), where the action denoted by the verbs cannot affect the objects. In short, GNOs must be affected by the action denoted by the verb.

4. SYNTACTIC REPRESENTATION

The aim of this chapter is to investigate whether GNOs are syntactically represented. The (syntactic) tests to show that a null argument is syntactically represented are related to control (sec. 4.1), binding (sec. 4.2) and secondary predication (sec. 4.3).

4.1. *Control*

GNOs are not syntactically represented cross-linguistically. In English, they cannot control. In (33), only the explicit object can control the interpretation of the null subject in the subordinate clause, as in (33a), while the same is not true with the null object, as in (33b) (Bach, 1979; Bresnan, 1982; Manzini, 1983; Lebeaux, 1984; Roberge, 1991; Landau, 1999; see Hornstein, 1999 and Wurmbrand, 2002 for discussion).

breaks one', which seems acceptable, versus *A veces los terremotos rompen* 'Sometimes earthquakes break one', which is completely out.

(33) *English*

- a. This leads people_i [PRO_i to conclude what follows].
- b. *This leads Ø_i [PRO_i to conclude what follows].

(Rizzi, 1986, p. 503)

Due to the contrast in (33), Bach (1979) proposed that, in object control structures, the object must be structurally represented. Because in English GNOs are not syntactically represented, the object must be made explicit in object control structures. This idea led Bach (1979) to establish the generalization in (34).

(34) *Bach's (1979) generalization*

In object control structures, the object must be structurally represented.

As opposed to English, Spanish GNOs can control the interpretation of null subjects. As shown in (35), the GNOs control the reference and interpretation of PROs. That is, GNOs follow the generalization of (34). (35b) means ‘ambition leads people to (those same people) make mistakes’. The same can be applied to (35a). That is, the GNO and PRO refer to the same entity (i.e., they coincide in their reference) (Lebeaux, 1984).

(35) *Control and GNOs*

- a. Esto lleva Ø_i a [PRO_i pensar lo siguiente].
this.NEUT lead.IND.PRS.3SG to think.IMP the.NEUT following.
‘This leads one to think the following.’
- b. La ambición conduce Ø_i a [PRO_i cometer errores].
the.FEM.SG ambition lead.IND.PRS.3SG to commit.IMP error.PL.
‘Ambition leads one to make mistakes.’

One (apparent) problem might have to do with the idea that the constructions in (35) are not transitive, but some kind of ‘complex’ predicates (i.e., *llevar a pensar* ‘lead one to think’ is a ‘complex’ predicate rather than a construction with two independent predicates, thus *llevar* ‘lead’ would lack an object). However, the data in (35) show transitive structures, because explicit counterparts of (35) can be found in (36).

(36) *Control and explicit objects*

- a. Esto lleva a la gente_i
this.NEUT lead.IND.PRS.3SG DOM the.FEM.SG people
a [PRO_i pensar lo siguiente].
to think.IMP the.NEUT following

‘This leads people to think the following.’

- b. La ambición conduce a la gente_i
 the.FEM.SG ambition lead.IND.PRS.3SG DOM the.FEM.SG people
 a [PRO_i cometer errores].
 to commit.INF error.PL
 ‘Ambition leads people to make mistakes.’

Another (apparent) counterexample to the generalization that Spanish GNOs are syntactically represented because they can participate in object control might be that GNOs in psychological predicates cannot control PRO. For instance, as opposed to (35), the sentences with psychological predicates in (37) do not display PRO.

(37) *Psychological predicates and GNOs*

- a. La música clásica relaja Ø.
 the.FEM.SG music classic.FEM.SG relax.IND.PRS.3SG
 ‘Classical music relaxes one.’
- b. Los sonidos de la naturaleza calman Ø.
 the.MASC.PL sound.PL of the.FEM.SG nature calm.IND.PRS.3PL
 ‘Sounds of nature soothe one.’

However, this restriction on PRO does not derive from the idea that GNOs are not syntactically represented in (37), but from the fact that psychological predicates (e.g., *preocupar* ‘worry’, *impresionar* ‘impress’, etc.) do not control PRO because they do not allow for non-finite sentences (i.e., they are lexically constrained in this sense).

4.2. Binding

As said above (see ch. 1), GNOs are not anaphoric in Spanish, which means that, in principle, they cannot be bound (Chomsky, 1981; Büring, 2005). Even in contexts where referents are introduced in preceding sentences, the GNO cannot refer to them, as in (38).

(38) *Anaphoric relations with GNOs*

- a. # [Virginia y su pareja]_i no pueden oír
 Virginia and her couple not AUX.IND.PRS.3PL hear.INF
 nada en el bar.
 nothing in the.masc.sg bar

- Tanto ruido deja \emptyset_i sordo.
 so-much.MASC.SG noise make.IND.PRS.3SG deaf.MASC.SG
 #‘Virginia and her partner cannot hear anything in the bar. So much noise deafens one.’
- b. #Juan sale siempre a correr.
 Juan go.IND.PRS.3SG always to run.INF
 Una rutina de ejercicio mantiene \emptyset_i sano.
 a.FEM.SG routine of exercise keep.IND.PRS.3SG healthy.MASC.SG
 #‘Juan always goes for a run. An exercise routine keeps one healthy.’

In (38a), the GNO cannot refer to Virginia or her partner, nor to Juan in (38b). Although Virginia and her partner are in (38a) within the set of individuals whom so much noise deafens, just as John is within the set of individuals whom a routine of exercise keeps healthy, the GNOs do not refer exclusively to Virginia and her partner in (38a) or to John in (38b) (i.e., co-reference is not involved). To refer to them, a definite clitic must be used. In (39a), *los* ‘them’ refers to Virginia and her partner, and in (39b), *lo* ‘him’ refers to Juan.

(39) *Anaphoric relations with clitics*

- a. Virginia y su pareja no pueden oír
 Virginia and her couple not AUX.IND.PRS.3PL hear.INF
 nada en el bar.
 nothing in the.masc.sg bar
 Tanto ruido los deja sordos.
 so-much noise them.MASC.PL make.IND.PRS.3SG deaf.MASC.PL
 ‘Virginia and her partner cannot hear anything in the bar. So much noise deafens them.’
- b. Juan sale siempre a correr.
 Juan go.IND.PRS.3SG always to run.INF
 Una rutina de ejercicio
 a.FEM.SGroutine of exercise
 lo mantiene sano.
 him.MASC.SG keep.IND.PRS.3SG healthy.MASC.SG
 ‘Juan always goes jogging. An exercise routine keeps him healthy.’

Also, the GNO does not seem to introduce a discourse referent, since, apart from cases of control, it cannot be retrieved by null subjects, as in (40a), or clitics, as in (40b). Null subjects and clitics obtain their interpretation from an antecedent, with which they are coreferent. However, in (40) they cannot refer to the GNOs:

they obtain their interpretation deictically, right from the context in which they are uttered (see Bosque 2015 for a detailed discussion). In this case, definite readings would be obtained, and no reference to GNOs would be made, contrary to what is intended.

(40) *Anaphoric relations with GNOs*

- a. ??Su inteligencia deja \emptyset_i atónito.
his/her intelligence leave.IND.PRS.3SG astonished.MASC.SG
 \emptyset_i Se quedaron boquiabiertos.
SE leave.IND.PRF.3PL dumbfounded.MASC.SG
??His/her intelligence leaves one astonished. They were dumbfounded.'
- b. ??Su inteligencia deja \emptyset_i atónito.
his/her intelligence leave.IND.PRS.3SG astonished.MASC.SG
Lo(s)_i dejó boquiabierto(s).
him(them).MASC.SG(PL) leave.IND.PRS.3SG speechless.MASC.SG(PL)
??His/her intelligence astounds. It left {him/them} dumbfounded.'

In principle, this suggests that GNOs do not introduce discourse referents (i.e., they are not phrases that can be referred to by an anaphoric expression). However, clitics can retrieve generic expressions. In (41), a clitic can refer to a generic antecedent such as *uno* 'one', which shares with GNOs a human, generic reading (see Gutiérrez-Rodríguez & Pérez Ocón, 2024 on Spanish *uno* 'one').

(41) *Anaphoric relations and generic uno 'one'*

- Su inteligencia deja a uno
his/her intelligence leave.IND.PRS.3SG DOM one.MASC.SG
sorprendido. iLo deja boquiabierto!
surprised.MASC.SG him.MASC.SG leave.IND.PRS.3SG speechless.MASC.SG
'His/her intelligence leaves one amazed. It leaves him speechless!'

Specifically, in (41) the clitic refers to the generic reading of the antecedent. In (41), the interpretation is 'leaves dumbfounded that same person who has been surprised by his intelligence'. Thus, the oddity of (40) is not due to the fact that GNOs do not introduce discourse referents, but because the clitic and the null subject cannot retrieve the generic interpretation of the antecedent, since they behave as pronouns (i.e., definite interpretations appear). As for null versions in (42), the arbitrary PROs must be linked in reference.

(42) *English*

PRO turning in one's tax forms on time requires PRO hiring an accountant.

(Authier, 1989, p. 48)

Although the reference of each PRO encompasses a generic set of individuals, the interpretation of the second PRO depends on that of the first PRO in (42). That is, (42) means 'for an individual x to turn in tax forms on time, it is necessary for that same individual x to hire an accountant' (see Lebeaux 1984; Authier 1989 for English and French data). That is, the PROs must coincide in their reference. This led Hernanz (1988) to think that PROs are variables linked to generic operators that make them have the same reference.

As far as Spanish GNOs are concerned, in (43) the reference of the second GNO is linked to the reference of the first one, as is the case with PRO in (42). That is, in (43) only the interpretation 'a medicine that heals a generic set of people is a medicine that benefits that same generic set of people' is obtained. These readings appear in equative structures (see Authier 1989 on this point), although it also appears with arbitrary null subjects and GNOs.

(43) *Equative structures and GNOs*

Una	medicina	que	sana \emptyset	es
a.FEM.SG	medicine	that	heal.IND.PRS.3SG	be.IND.PRS.3SG
una	medicina	que	beneficia \emptyset .	
a.FEM.SG	medicine	that	benefit.IND.PRS.3SG	

'A medicine that heals one is a medicine that benefits one.'

In (43), the interpretation by which a medicine that heals any person is a medicine that benefits another person is inadequate. This is because both GNOs have to be linked in some interpretative sense. That is, the readings of both GNOs must be coincident. To illustrate this more clearly, see the contrast in (44).

(44) *Equative structures and GNOs*

a. La	música	clásica	que	entretiene \emptyset
the.FEM.SG	music	classic.FEM.SG	that	entertain.IND.PRS.3SG
es	la	que	relaja \emptyset .	

be.IND.PRS.3SG the.FEM.SG that relax.IND.PRS.3SG

'The classical music that entertains one is the one that relaxes one.'

- b. #La música clásica que relaja \emptyset
 the.FEM.SG music classic.FEM.SG that relax.IND.PRS.3SG
 es la que pone nervioso \emptyset .
 be.IND.PRS.3SG the.FEM.SG that make.IND.PRS.3SG nervous.MASC.SG
- #‘The classical music that relaxes one is the one that makes one nervous.’

In (44a), the classical music that entertains any person is the music that relaxes the same person. This is the only reading available. If each GNO referred to different individuals, (44b) would not be anomalous, since the people who are relaxed by classical music do not coincide with the people who are made nervous by classical music. However, since both GNOs must be referentially linked, it is not clear why in (44b) the classical music that relaxes any person is the one that makes that same person nervous. This is a contradiction.

It might be thought that the oddity of (44b) is related to the contradiction caused by the predicates *relajar* ‘relax’ and *poner nervioso* ‘make nervous’, which are related by a copulative structure. However, if this were so, one would expect that a sentence with the same predicates, but with explicit objects, would be anomalous. However, (45), where objects referring to different individuals have been introduced, and the predicates in (44b) have been maintained, is grammatical. Thus, the oddity of (44b) is not due to a contradiction by the predicates, but to the fact that the GNOs must display the same interpretation.

(45) *Equative structures and explicit objects*

- La música clásica que relaja
 the.FEM.SG music classic.FEM.SG that relax.IND.PRS.3SG
 a Juan es la que pone
 DOM Juan be.IND.PRS.3SG the.FEM.SG that make.IND.PRS.3SG
 nervioso a José.
 nervous.MASC.SG DOM José

‘The classical music that relaxes Juan is the one that makes José nervous.’

As for binding with reflexive pronouns, in Spanish the clitic *te* ‘(to) you’ and the reflexive pronoun *contigo mismo* ‘with yourself’ can be linked, as in (46a), as well as the generic pronoun *uno* ‘one’ and the reflexive pronoun *consigo mismo* ‘with oneself’, as in (46b). This is so because, in the former case, both elements ‘agree’ in second person, and, in the latter case, they ‘agree’ in third person.

(46) *Binding and pronouns*

- a. La buena música te reconcilia
 the.FEM.SG good.FEM.SG music you reconcile.IND.PRS.3SG
 contigo mismo.
 with-you yourself.MASC.SG

‘Good music reconciles you with yourself.’

- b. La buena música reconcilia
 the.FEM.SG good.FEM.SG music reconcile.IND.PRS.3SG
 a uno consigo mismo.
 DOM one with-him himself.MASC.SG

‘Good music reconciles one with oneself.’

This anaphoric relation is only possible with explicit objects such as those in (46). In contrast, GNOs cannot function as antecedents of these same reflexive pronouns, as in (47). This is so because, as opposed to the cases in (46), the GNO and the pronoun do not share person features.

(47) *Binding, pronouns and GNOs*

- a. ??La buena música reconcilia Ø
 the.FEM.SG good.FEM.SG music reconcile.IND.PRS.3SG
 contigo mismo.
 with-you yourself.MASC.SG
- b. *La buena música reconcilia Ø
 the.FEM.SG good.FEM.SG music reconcile.IND.PRS.3SG
 consigo mismo.
 with-him himself.MASC.SG

These sentences are only grammatical with *uno mismo* ‘oneself’. In (48), the reference of *uno mismo* ‘oneself’ coincides with that of the GNO (i.e., the interpretation of the proposition in (48) is ‘good music reconciles one to oneself’). As the GNO and the pronoun *uno mismo* ‘oneself’ display third person, the sentence is well-formed. In fact, GNOs cannot be interpreted as a first or second person.

(48) *Anaphoric relation between uno and GNOs*

- La buena música reconcilia Ø
 the.FEM.SG good.FEM.SG music reconcile.IND.PRS.3SG
 con uno mismo.

with one.MASC.SG self.MASC.SG
 'Good music reconciles one with oneself.'

This evidence indicates that GNOs introduce discourse referents. This leads us to think that GNOs are not only semantically interpreted, but also syntactically represented.

4.3. Secondary predication

If a GNO is visible as a subject of predication, it is syntactically represented. If it were not, the sentence would be ungrammatical, since the secondary predicate could not be predicated of any syntactic element (see Gumié, 2005, 2007; Landau, 2010 on secondary predication). As explained above for Italian and French (see ch. 2), Spanish GNOs are also visible as subject of secondary predication, as in (49).

(49) *Secondary predication and GNOs*

- a. El dinero no hace feliz.
 the.MASC.SG money not make.IND.PRS.3SG happy
 'Money does not make one happy.'
- b. Su inteligencia deja atónito.
 his/her intelligence leave.IND.PRS.3SG astonished.MASC.SG
 'His/her intelligence leaves one astonished.'

In (49), the secondary predicate must be predicated of the GNO. The reading of (49a) is 'money does not make anyone happy'. The same can be applied in (49b). Similarly, with explicit objects, as in (50a), the secondary predicate can only be predicated of the object. (49) and (50) behave the same way, but they receive different interpretations: the null objects in (49) receive generic readings, while the explicit objects in (50) receive particularizing ones.

(50) *Secondary predication and explicit objects*

- a. El dinero no hace feliz a Pedro.
 the.MASC.SG money not make.IND.PRS.3SG happy DOM Peter
 'Money does not make Peter happy.'
- b. Su inteligencia deja atónito a Pedro.
 his/her intelligence leave.IND.PRS.3SG astonished.MASC.SG DOM Peter
 '{His/her} intelligence leaves Peter astonished.'

However, in (51), the secondary predicate cannot be predicated of the GNO, contrary to (49). The proposition in (51) can only be interpreted in such a way that the subject functions as the subject of the secondary predicate, but not the GNO.

(51) *Secondary predication and GNOs*

El	doctor	ausculta	Ø	desnudo.
the.MASC.SG	doctor.MASC.SG	auscultate.IND.PRS.3SG	naked.MASC.SG	
'The doctor auscultates naked people.'				

- *Subject predication*: 'The doctor auscultates people naked.'

- *Object predication*: # 'The doctor auscultates naked people.'

Interestingly, if the object is explicit, (51) is ambiguous between an interpretation in which the secondary predicate is predicated of the subject and one in which the secondary predicate is predicated of the object, as in (52).

(52) *Secondary predication and explicit objects*

El	doctor	ausculta	a	Pedro	desnudo.
the.MASC.SG	doctor.MASC.SG	auscultate.IND.PRS.3SG	DOM	Peter	naked.MASC.SG
'The doctor auscultates Peter naked.'					

- *Subject predication*: 'The doctor auscultates Peter naked.'

- *Object predication*: 'The doctor auscultates Peter when Peter is naked.'

Two questions arise from these contrasts. The first question is why the secondary predicate is predicated of the GNO in (49), but not in (51). This question leads to the second question: why ambiguity between two readings is possible in (52), but not in (51).

The answer has to do with syntactic structure. While (49) shows causative structures, by which the object undergoes a change of state (whose resulting state is denoted by the secondary predicate), in (51) there is no causative structure. Thus, the small clause composed of GNOs and secondary predicates in (51) is not argumental, as opposed to those in (49). In causative structures, the only available reading is the one in which the secondary predicate is predicated of the object (whether explicit or implicit). Hence, the examples of (49)–(50) are not ambiguous between subject and object predication, and in (51) the secondary predicate can only be predicated of the subject, since they are not causative structures. This restriction explains why the secondary predicates of (49) are predicated of the GNO, but cannot do the same in (52). Thus, the absence of object predication in

(51) is not related to the fact that there is no syntactic entity projected in object position (i.e., GNOs in Spanish are not only semantically interpreted, but also syntactically represented).

Secondary predicates also reveal the features of GNOs in Spanish. Specifically, secondary predicates usually display masculine gender, as in (53). The fact that the default gender of GNO is masculine is coherent with the generic reading of the object and sentence. Since the property designated by the subject must apply to the arbitrary persons designated by the GNO, the unmarked gender appears, which in Spanish is masculine. This indicates that the gender of GNOs coincides with the default gender in Spanish.

(53) *Gender and GNOs*

Tanto	ruido	deja Ø	sordo.
so-much	noise	leave.IND.PRS.3SG	deaf.MASC.SG
‘So much noise deafens one.’			

However, in contexts where the generalization applies exclusively to women, the GNO may display feminine gender. In a bachelor’s party where only women are present, it might be the case that (54), where the adjective shows feminine gender (i.e., marked gender), is uttered. That GNOs show feminine gender does not prevent either the generic reading of the sentence or the object to be possible. In fact, the characteristic ‘make deaf’ of the noise in (54a) still applies to any individual hearing such a noise. The difference is that the scope is restricted to all *x* that meet the condition of being female.

(54) *Gender variation and GNOs*

Tanto	ruido	deja Ø	sorda.
so-much	noise	leave.IND.PRS.3SG	deaf.FEM.SG
‘So much noise leaves one deaf.’			

Secondary predication also shows that GNOs display number. In Spanish, GNOs cannot show plural number, as in (55). Even if the interlocutors were in a situation where only men were present, it could not be the case that the utterances of (55) are uttered. The number features must therefore correspond to the default forms, that is, to the singular forms.

(55) *Number variation and GNOs*

- a. *El dinero no hace \emptyset felices.
 the.MASC.SG money not make.IND.PRS.3SG happy.PL
 b. *Su inteligencia deja \emptyset atónitos.
 his/her intelligence leave.IND.PRS.3SG astonished.MASC.PL

These data contrast with their explicit counterparts, where the object agrees in plural with the secondary predicate, as in (56).

(56) *Number variation and explicit objects*

- Su inteligencia deja a todos atónitos.
 his/her intelligence leave.IND.PRS.3SG DOM all.MASC.PL astonished.MASC.PL
 ‘{His/her} intelligence leaves everyone astonished.’

Secondary predication also shows that GNOs display a syntactically projected D. Secondary predication readings are only obtained with definite arguments (Demonte & Masullo, 1999). According to Landau (2010), null arguments that are visible as the subject of secondary predication have syntactically projected D. This is consistent with the idea that bare nouns lack D, given that they are syntactically projected NPs (see Dobrovie-Sorin & Laca, 2003 for a detailed discussion). In contrast, Spanish GNOs show a syntactically represented D feature (i.e., they are projected as full DPs). Note that, from a theoretical point of view, it makes sense to make this move, since only DPs (not NPs) can receive generic interpretations in Spanish (see Dobrovie-Sorin & Laca, 2003 on this point).

5. A PRELIMINARY ANALYSIS

In this chapter it is proposed a tentative analysis that attempts to explain the above-mentioned grammatical properties of GNOs in Spanish. The analysis, in the spirit of Authier (1989, 1992), relies on variables.

As explained above (see sec. 2), quantified phrases may interact to give rise to scope ambiguities. (57) is ambiguous in two ways: as shown below, it can either mean ‘in this paradise, there is something that relaxes all the new visitors’ or ‘for all the new visitors, there is something that relaxes each one’. Following May (1977) and Authier (1989), we assign the sentence in (57) the representations in (58) at LF. Thus, in the first reading the existential quantifier is ‘above’ the universal one, as shown in (58a), while in the second reading the universal quantifier is ‘above’ the existential one, as shown in (58b).

(57) *Scope ambiguities*

- En este paraíso, alguna cosa relaja
 in this.MASC.SG paradise some.FEM.SG think relax.IND.PRS.3SG
 a todos los nuevos visitantes.
 DOM all.MASC.PL the.MASC.PL new.MASC.PL visitor.PL
 'In this paradise, something relaxes all the new visitors.'
 - 1st reading: $\exists x, x$ is a thing, such as x relaxes $\forall y, y$ is a person.
 - 2nd reading: $\forall y, y$ is a person, $\exists x, x$ is a thing, such as x relaxes y .

(58) *Representation at LF*

- a. 1st reading
 ... [alguna cosa]_i [todos los nuevos visitantes]_j [x_i relaja x_j]
 b. 2nd reading
 ... [todos los nuevos visitantes]_j [alguna cosa]_i [x_i relaja x_j]

Consider next (59), which contains a GNO and an existentially quantified phrase. This example is (again) two ways ambiguous: as shown below, it can either mean 'in this paradise, there is something that relaxes people' (i.e., the existential quantifier is 'above' the universal one, giving rise to the first reading) or 'for all people visiting this paradise, there is something that relaxes each one' (i.e., the universal quantifier is 'above' the existential one, giving rise to the second reading).

(59) *Scope ambiguities*

- En este paraíso, alguna cosa relaja Ø.
 in this.MASC.SG paradise some.FEM.SG think relax.IND.PRS.3SG
 'In this paradise, something relaxes one.'
 - 1st reading: $\exists x, x$ is a thing, such as x relaxes $\forall y, y$ is a person.
 - 2nd reading: $\forall y, y$ is a person, $\exists x, x$ is a thing, such as x relaxes y .

Note that the readings in (59) nicely parallel those in (57). This is so because the GNO in (59) interacts with the existentially quantified subject in the same fashion as the universal quantifier phrase in (57). Thus, assuming that GNOs are variables bound by an operator explains the interpretive similarity between (57) and (59). The remaining question is what kind of operator binds the variable. In answering this question, Authier (1989) makes use of unselective binding. Consider the examples in (60).

(60) *Unselective binders and GNOs*

- a. Normalmente, este paraíso relaja Ø.
 normally this.MASC.SG paradise relax.IND.PRS.3SG
 'Normally, this paradise relaxes one.'
- b. Este paraíso jamás relaja Ø.
 this.MASC.SG paradise never relax.IND.PRS.3SG
 'This paradise never relaxes one.'

The GNOs do not seem to have quantificational force of their own. Rather, they reflect the quantificational force of the quantified phrase in the sentence. Thus, the examples in (60) seem amenable to the paraphrases in (61). Specifically, as the quantified phrase in (60a) is *normalmente* 'typically', the GNO is interpreted as 'people', in general, as shown in (61a); conversely, as the quantified phrase in (60b) is *jamás* 'never', the GNO is interpreted as 'anyone', as shown in (61b).

(61) *Unselective binders and explicit objects*

- a. Este paraíso relaja a mucha gente.
 this.MASC.SG paradise relax.IND.PRS.3SG DOM much.FEM.SG people
 'This paradise relaxes many people.'
- b. Este paraíso no relaja a nadie.
 this.MASC.SG paradise not relax.IND.PRS.3SG DOM anyone
 'This paradise does not relax anyone.'

Crucially, the GNO in the version of (60a) without the quantifier phrase (e.g., *Este paraíso relaja Ø* 'this paradise relaxes') seems to have the force of universal quantification too, but there is no overt binder for the variable in this case. As explained above (see sec. 3), the distribution of GNOs is restricted by genericity. Thus, in these cases there is a non-overt adverb of quantification which is induced «by making reference to the properties of Infl and in particular to tense» (Authier, 1989, p. 56). This non-overt adverb is close in interpretation to *normalmente* 'normally' in Spanish. In other words, when there is no overt quantified phrase like *normalmente* 'typically', *never*, etc., the sentence displays a default interpretation with *normalmente* 'typically'.

These facts suggest that GNOs are generated as free variables which get bound at LF (Lewis, 1975). Thus, in line with Authier (1989, 1992), the analysis relies on representing the GNOs as free variables, as in (62), where *e* represents the variable.

The operator can be explicit, as *jamás* ‘never’ in (62a), or implicit, as in (62b) (see Lewis, 1975 on this point).⁶

- (62) Tentative analysis
- a. $\text{Jamás}_i \left[\text{VP} \left[\text{V relaja} \right] \left[\text{DP} e_i \right] \right]$
 - b. $\text{OP}_i \left[\text{VP} \left[\text{V relaja} \right] \left[\text{DP} e_i \right] \right]$

Thus, analyzing GNOs as free variables whose interpretation depends on the operator of the sentence correctly explains the above-mentioned properties of GNOs. However, this is just a preliminary analysis: further research is needed in order to confirm its shortcomings and predictions.

6. CONCLUDING REMARKS

This paper has explored the constraints on the occurrence of non-anaphoric null objects, their syntactic representation and their grammatical properties in Spanish (e.g., *Esta película sorprende Ø* ‘this film surprises’).

Regarding the constraints on GNOs in Spanish, they must be non-anaphoric (i.e., they are interpreted without an antecedent). Additionally, they must occur in generic sentences and receive generic interpretations. That is why they cannot occur with imperfective tenses (e.g., **Esta película sorprendió Ø* ‘this film surprised’), but are natural with imperfective ones (e.g., *Esta película sorprendía Ø* ‘this film surprised’). Furthermore, GNOs must receive a human reading. This restriction explains why GNOs are impossible with predicates that only allow inanimate objects, such as *haber* ‘there is’ or *fabricar* ‘manufacture’ (e.g., **En esta fábrica fabrican Ø rápido* ‘in this factory they manufacture quickly’; **En esta tienda había Ø* ‘in this there was someone’). In other words, their licensing is lexically restricted. Additionally, GNOs must be affected by the event denoted

6. As correctly pointed out by an anonymous reviewer, while the analysis favors a variable-based interpretation, some properties of GNOs (such as syntactic projection, participation in predication, and interaction with reflexives) resemble those of pronominal expressions. In fact, as pointed out above, secondary predicates are only possible with full DPs. Since GNOs can be combined with secondary predicates, GNOs must be projected as full DPs (as opposed to indefinite null objects, which behave as bare NPs; see Martínez-García, 2023, 2025 for further details). Provided they are null DPs, and that they can interact with reflexives, it seems reasonable to propose that they behave as null pronominals (i.e., *pro*; Chomsky 1981). This means that, from a semantic point of view, GNOs seem to behave as variables, but from a syntactic point of view, they seem to behave as empty pronominals. Although this is a very interesting issue to address, due to the fact that we are mostly concerned about how Spanish GNOs could relate to French and Italian GNOs and to limitations of space, the potential hybrid status of GNOs (as DPs with variable-like content but pronominal syntax) should be explored more thoroughly in future research.

by the verb. Specifically, they must receive a thematic role different from Theme. This explains why they cannot occur with predicates that do not denote any affectedness of the object (e.g., **Juan admira Ø* 'Juan admires'), but are natural with causative predicates, which entail affected objects (e.g., *Este sonido deja Ø sordo* 'this noise deafens').

Once the conditions of occurrence and interpretation of GNOs have been described, their syntactic structure has been explored. GNOs are not only semantically interpreted, but also syntactically represented, since they are sensitive to syntactic operations such as control, binding and secondary predication. These syntactic operations show that GNOs display gender features (typically masculine gender, but also feminine under specific situations), and number features, specifically singular (i.e., the non-marked form in Spanish). Additionally, GNOs show D features, given that they are visible as subject of secondary predication, contrary to NPs. Specifically, they are syntactically a set of D and φ -features (gender and number). Thus, it has been proposed in this paper that GNOs are full DPs in Spanish. Note that, from a theoretical point of view, it makes sense to make this move, since only DPs (not NPs) can receive generic interpretations in Spanish.

Finally, based on previous accounts on GNOs in Italian and French, a tentative analysis for GNOs in Spanish has been proposed. Specifically, the analysis relies on representing the GNOs as free variables. Thus, GNOs are generated as free variables bound in LF by an (explicit or implicit) operator. This analysis explains the different interpretations that GNOs can receive in Spanish.

Nonetheless, there are some remaining issues worth addressing in future research. First, it seems to be true that most languages display some kind of GNOs, but it is unclear whether their features are the same regarding animacy, genericity and affectedness, and whether the proposed analysis could be extended cross-linguistically. It seems that the analysis is predictive enough for Italian, French and Spanish GNOs, but further research is needed regarding other (Romance) languages (e.g., it would be interesting to have a look at Portuguese GNOs, which have never been studied before, as far as I am concerned). If GNOs behave the same way cross-linguistically, we would discover a very interesting pattern across languages that would suggest core cross-linguistic grammatical features. Second, even though Italian, French and Spanish GNOs have now received attention in the literature, it would be interesting to study how the variety affects the licensing of GNOs in each language. Regarding Spanish, it is expected that the use of GNOs is dependent on the type of predicates each variety uses, but further research is needed in this regard.

Third, as explained above, while the analysis favors a variable-based interpretation, some properties of GNOs (such as syntactic projection, participation in predication, and interaction with reflexives) resemble those of pronominal expressions. This means that, from a semantic point of view, GNOs seem to behave as variables, but from a syntactic point of view, they seem to behave as empty pronominals. Although this is a very interesting issue to address, the potential hybrid status of GNOs should be explored more thoroughly in future research.

In sum, this paper has pointed out the existence of Spanish generic object drop, and has identified the grammatical properties of GNOs and the conditions that must be met to allow generic object drop. Future research might shed light on new data that strengthen the predictive capability of the analysis.

ACKNOWLEDGEMENTS

Related versions of this paper were presented at *Kolloquium für Examenskandidat*innen und Doktorand*innen* (Goethe-Universität, November 15, 2022), at the panel «La elipsis y la arquitectura del lenguaje: ¿Cuándo puede elidirse qué y por qué?», within *LII Simposio de la SEL* (CCHS-CSIC, January 25, 2024), at *Seminario de Lingüística* (Universidad Complutense de Madrid, February 5, 2024), and at *LIII Simposio de la SEL* (Universidad de Salamanca, January 20, 2025). I thank the audience for the valuable suggestions. I am also grateful to the two anonymous reviewers of *RSEL*, whose insightful comments have greatly improved this paper. All mistakes and shortcomings in the work are my own.

FUNDING

This research was funded by the Spanish Ministerio de Universidades, under the auspices of FPU20/04298, EST23/00059 and EST24/00098, by the Spanish Agencia Estatal de Investigación, with funded project PID2021-123617NB-C3, and by Departamento de Lengua Española y Teoría de la Literatura (Universidad Complutense de Madrid).

REFERENCES

- Authier, J.-M. (1989). Arbitrary null objects and unselective binding. In O. A. Jaeggli & K. Safir (Eds.), *The Null Subject Parameter* (pp. 45–67). Springer Netherlands.
doi: <https://doi.org/10.1007/978-94-009-2540-3>
- Authier, J.-M. (1992). Arbitrary null object languages in a parametric theory of linguistic variation. *Anuario del Seminario de Filología Vasca «Julio de Urquijo»*, 27, 37–52.
- Bach, E. (1979). Control in Montague grammar. *Linguistic Inquiry*, 10, 515–531.

- Barbosa, P. (2019). *Pro as a minimal nP: Toward a unified approach to pro-drop*. *Linguistic Inquiry*, 50, 487–526. doi: https://doi.org/10.1162/ling_a_00312
- Barbosa, P. (2024). Null objects, null nominal anaphora and antilogophoricity. *Probus*, 36, 251–281. doi: <https://doi.org/10.1515/probus-2024-2012>
- Beavers, J. (2011). On affectedness. *Natural Language & Linguistic Theory*, 29, 335–370. doi: <https://doi.org/10.1007/s11049-011-9124-6>
- Bosque, I., & Gutiérrez-Rexach, J. (2009). *Fundamentos de sintaxis formal*. AKAL.
- Bresnan, J. (1982). Control and complementation. *Linguistic Inquiry*, 13, 343–434.
- Brucart, J. M.^a (1999). La elipsis. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española*, Vol. 2 (pp. 2787–2863). Espasa-Calpe.
- Büring, D. (2005). *Binding theory*. Cambridge University Press.
doi: <https://doi.org/10.1017/CBO9780511802669>
- Campos, H. (1986). Indefinite object drop. *Linguistic Inquiry*, 17, 354–359.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, 339–405. doi: <https://doi.org/10.1023/A:1008324218506>
- Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications.
- Clements, J. C. (2006). Null direct objects in Spanish. In J. C. Clements & J. Yoon (Eds.), *Functional approaches to Spanish syntax* (pp. 134–150). Palgrave Macmillan.
doi: https://doi.org/10.1057/9780230522688_6
- Cummins, S., & Roberge, Y. (2004). Null objects in French and English. In J. Auger, J. C. Clements & B. Vance (Eds.), *Contemporary Approaches to Romance Linguistics: Selected Papers from the 33rd Linguistic Symposium on Romance Languages* (pp. 121–138). Bloomington. doi: <https://doi.org/10.1075/cilt.258.07cum>
- Demonte, V., & Masullo, P. J. (1999). La predicación: los complementos predicativos. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española*, Vol. 2 (pp. 2461–2523). Espasa-Calpe.
- Dobrovie-Sorin, C. (1997). Existential arguments and existential predicates. In A. Lawson (Ed.), *Proceedings of Semantics and Linguistic Theory VII* (pp. 117–134). CLC Publications. doi: <https://doi.org/10.3765/salt.v7i0.2796>
- Dobrovie-Sorin, C., & Laca, B. (2003). Les noms sans déterminant dans les langues romanes. In D. Godard (Ed.), *Les langues romanes. Problèmes de la phrase simple* (pp. 235–281). Éditions du CNRS.
- Dowty, D. R. (1991). Thematic proto-roles and argument selection. *Language*, 67, 547–619. doi: <https://doi.org/10.1353/lan.1991.0021>
- Dvořák, V. (2017). *Generic and indefinite null objects*. PhD dissertation. Rutgers University.
- Fábregas, A. (2023). Ellipsis in Spanish. *Borealis*, 12, 1–134.
doi: <https://doi.org/10.7557/1.12.2.7228>
- Gumiel Molina, S. (2005). *Los complementos predicativos*. Arco/Libros.
- Gumiel Molina, S. (2007). *Estructura argumental y predicación secundaria*. Servicio de Publicaciones de la Universidad de Alcalá.
- Gutiérrez-Rodríguez, E., & Pérez Ocón, P. (2024). Inclusive readings of the Spanish prounoun ‘uno’ (‘One’): Grammatical and pragmatic properties. *Revista Española de Lingüística*, 54, 111–140. doi: <https://doi.org/10.31810/rsel.54.2.3>

- Hernanz, M. L. (1988). En torno a los sujetos arbitrarios de segunda persona del singular. In V. Demonte & B. Garza (Eds.), *Estudios lingüísticos de España y de México* (pp. 151–178). El Colegio de México-UNAM. doi: <https://doi.org/10.2307/j.ctv43vs5t.9>
- Hornstein, N. (1999). Movement and control. *Linguistic Inquiry*, 30, 69–96. doi: <https://doi.org/10.1162/002438999553968>
- Jaeggli, O. (1982). *Topics in Romance syntax*. Foris Publications.
- Jaeggli, O. (1986). Arbitrary plural pronominals. *Natural Language & Linguistic Theory*, 4, 43–76. doi: <https://doi.org/10.1007/BF00136264>
- Kim, S. (1991). A uniform analysis of arbitrary null subjects and objects. *Proceedings of WCCFL*, 10, 283–296.
- Krifka, M. et al. (1995). Genericity: an introduction. In G. N. Carlson & F. J. Pelletier (Eds.), *The generic book* (pp. 1–124). The University of Chicago Press.
- Laca, B. (2013). Spanish bare plurals and topicalization. In J. Kabatek & A. Wall (Eds.), *New perspectives on bare noun phrases in Romance languages and beyond* (pp. 95–120). John Benjamins. doi: <https://doi.org/10.1075/slcs.141.04lac>
- Landau, I. (1999). *Elements of control*. PhD dissertation. MIT.
- Landau, I. (2010). The explicit syntax of implicit arguments. *Linguistic Inquiry*, 41, 357–388. doi: https://doi.org/10.1162/LING_a_00001
- Lebeaux, D. (1984). Anaphoric binding and the definition of PRO. *North East Linguistic Society*, 14, 253–274.
- Leonetti, M. (1990). *El artículo y la referencia*. Taurus.
- Leonetti, M. (1999). El artículo. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española*, Vol. 1 (pp. 787–890). Espasa-Calpe.
- Lewis, D. (1975). Adverbs of quantification. In E. L. Keenan (Ed.), *Formal semantics of natural languages* (pp. 3–15). Cambridge University Press. doi: <https://doi.org/10.1017/CBO9780511897696.003>
- Manzini M. R. (1983). On control and control theory. *Linguistic Inquiry*, 14, 421–446.
- Martínez-García, C. (2023). Indefinite readings of referential null subjects and null objects in Spanish. *Borealis*, 12, 135–159. <https://doi.org/10.7557/1.12.2.7259>
- Martínez-García, C., & Romero Heredero, D. (2024). Marcado diferencial de objeto y objetos tácitos: Hacia una comparación en español. In S. Gumié-Molina, I. Bosque & C. Sánchez López (Eds.), *Buceando entre palabras. Inmersiones en gramática de la mano de Manuel Leonetti* (pp. 261–274). Editorial de la Universidad de Alcalá.
- Martínez-García, C. (2025). Indefinite argument drop in Spanish. In K. Colleluori (Ed.), *Proceedings of the thirty-sixth WECOL* (pp. 114–123). California State University, Fresno.
- Masullo, P. J. (2015). La interfaz sintaxis-semántica/pragmática: sobre la arbitrariedad de *pro_{arb}*. *SAGA, Revista de Letras*, 4, 200–230.
- May, R. C. (1977). *The grammar of quantification*. PhD dissertation. MIT.
- Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*, 17, 501–557.
- Roberge, Y. (1991). On the recoverability of null objects. In D. Wanner & D. A. Kibbee (Eds.), *New analyses in Romance linguistics* (pp. 299–312). John Benjamins. doi: <https://doi.org/10.1075/cilt.69.22rob>
- Suñer, M. (1983). *Pro_{arb}*. *Linguistic Inquiry*, 14, 188–191.

- Verdecchia, M. (2022). Object drop in Spanish is not island-sensitive. *Journal of Linguistics*, 58, 901-906. doi: <https://doi.org/10.1017/S002226722000202>
- Wurmbrand, S. (2002). Syntactic vs. semantic control. In C. J.-W. Zwart & W. Abraham (Eds.), *Studies in comparative Germanic syntax. Proceedings from the 15th Workshop on Comparative Germanic Syntax* (pp. 93–127). John Benjamins. doi: <https://doi.org/10.1075/la.53.07wur>



DIVERSIDAD Y SOFISTICACIÓN DE COLOCACIONES EN LA PRODUCCIÓN ESCRITA DE APRENDICES Y HABLANTES NATIVOS DE ESPAÑOL

Rocío CUBEROS-VICENTE¹
Universidad Complutense de Madrid

Elisa ROSADO²
Universidad de Barcelona

Resumen

El desarrollo del lenguaje formulaico en segundas lenguas (L₂) ha sido analizado en estudios que señalan el uso de colocaciones como indicador fiable de competencia lingüística. Sin embargo, las colocaciones han recibido poca atención en la investigación sobre complejidad lingüística. Este estudio pretende contribuir a dicha investigación examinando la complejidad fraseológica, constructo que refiere a la diversidad y la sofisticación de las colocaciones utilizadas en la producción escrita de aprendices de nivel intermedio y avanzado y hablantes nativos de español. Se analizan las colocaciones verbo-sustantivo y sustantivo-adjetivo en dos tipos de texto: narrativo y expositivo. Los datos provienen de dos tareas de producción escrita espontánea a partir de un estímulo visual. Los resultados indican que la complejidad fraseológica aumenta con el nivel de competencia en ambos tipos de texto, pero su desarrollo varía según la estructura sintáctica. Se discuten las implicaciones de estos hallazgos para la enseñanza de L₂ y se sugieren orientaciones para futuros estudios.

Palabras clave: español lengua segunda; colocaciones; complejidad fraseológica; diversidad; sofisticación.

DIVERSITY AND SOPHISTICATION OF COLLOCATIONS IN WRITTEN PRODUCTION OF LEARNERS AND NATIVE SPEAKERS OF SPANISH

Abstract

The development of formulaic language in second language (L₂) has been analyzed in studies that point to the use of collocations as a reliable indicator of linguistic proficiency.

1. marcubert@ucm.es;  <https://orcid.org/0000-0002-9089-5092>
2. erosado@ub.edu;  <https://orcid.org/0000-0002-1459-284X>

However, collocations have received little attention in research on linguistic complexity. The present study aims to contribute to this line of research by analyzing phraseological complexity, a construct that refers to the diversity and sophistication of collocations used in written production of intermediate and advanced learners and native speakers of Spanish. Verb-noun and noun-adjective collocations are analyzed in narrative and expository texts. The data comes from two spontaneous writing production tasks based on a visual stimulus. The results show that phraseological complexity increases with proficiency, with equivalent patterns across text types, but the developmental pattern did vary according to the syntactic structure of collocations. The implications of these findings for L2 teaching are discussed and directions for future research are suggested.

Keywords: Spanish as a second language; collocations; phraseological complexity; collocational diversity; collocational sophistication.

RECIBIDO: 21/02/2025

ACEPTADO: 14/11/2025

1. INTRODUCCIÓN

El lenguaje formulaico desempeña un papel fundamental en el desarrollo del lenguaje nativo y no nativo, en tanto que el dominio de determinadas expresiones prefabricadas puede facilitar el aprendizaje de lenguas primeras (L1) y segundas (L2) (Ellis et al., 2015; Hyland, 2008; Schmitt, 2004). La reflexión metalingüística sobre la composición de dichas expresiones puede suponer para el aprendiz una valiosa estrategia de análisis que, en última instancia, contribuya al aprendizaje de la L2 (Ellis et al., 2015). Además, el conocimiento y uso de estas expresiones aporta fluidez y naturalidad al discurso, por cuanto constituyen opciones lingüísticas preferidas frente a otras alternativas gramaticalmente posibles, pero menos frecuentes (Kecske, 2007). Estudios previos sobre la relación entre la calidad de los textos y sus características léxicas (Crossley et al., 2012; Paquot, 2018) coinciden en señalar la necesidad de considerar el uso de unidades léxicas multipalabra para alcanzar una correcta valoración de la calidad de los textos producidos por aprendices de L2. Por ejemplo, en Cuberos (2019) se demostró cómo el uso de colocaciones por parte de hablantes nativos y no nativos de español está asociado no solo a niveles de competencia más altos, sino también a la evaluación de la calidad textual, por cuanto su uso contribuyó a explicar variaciones en las calificaciones asignadas a los textos por parte de jueces expertos.

Dentro del estudio del lenguaje formulaico en el ámbito del español L2, las colocaciones léxicas han recibido especial atención en los últimos años (véase, por

ejemplo, García-Salido et al., 2019; Pérez-Serrano y Mavrou, 2022), principalmente, porque su dominio representa un desafío para los aprendices de L2 (Fan, 2009; Nesselhauf, 2004). Los estudios relacionados con el uso de colocaciones léxicas en español L2 se han centrado sobre todo en caracterizar el repertorio de colocaciones en las producciones de aprendices en contraste con el equivalente nativo y, en menor medida, en compararlo en distintos niveles de competencia. Todos ellos coinciden en señalar el carácter deficitario del repertorio de colocaciones empleado en español L2, de forma que los aprendices no solo suelen emplear colocaciones erróneas que se desvían de la equivalente nativa sino que, además, tienden a sobreutilizar determinadas colocaciones, ya sea por su alta frecuencia en la lengua meta o por la similitud con la L1 (García-Salido, 2016; García-Salido y García, 2017; García-Salido et al., 2019). Con respecto a su uso según el nivel de competencia, se observa un aumento progresivo en la cantidad, diversidad y corrección de las colocaciones a medida que aumenta el grado de dominio lingüístico (Cuberos, 2019; Cuberos et al., 2024; García-Salido et al., 2019). Sin embargo, incluso en niveles avanzados, el repertorio de colocaciones empleado sigue sin ser equivalente al de los hablantes nativos (García-Salido et al., 2019). Estos resultados coinciden con los obtenidos sobre el uso de colocaciones por parte de aprendices de otras lenguas como el inglés (Durrant y Schmitt, 2009), el francés (Vandeweerd et al., 2021) o el alemán (Boone y Eyckmans, 2023).

Este estudio pretende contribuir a la caracterización del repertorio de colocaciones del español nativo y no nativo en dos corpus equivalentes. Para ello, examinamos dicho repertorio atendiendo al concepto de *complejidad fraseológica*, que alude a dos propiedades del uso que se hace de las colocaciones, a saber, la *diversidad colocacional* y la *sofisticación colocacional* (Paquot, 2019). Mientras que la *diversidad colocacional* refiere a la variedad de colocaciones utilizadas en un texto, el concepto de *sofisticación colocacional* identifica el uso de colocaciones menos comunes y propias de niveles avanzados. En español como L2, estas propiedades han sido analizadas en el conjunto de colocaciones producidas por los aprendices de modo general (véase, por ejemplo, García-Salido et al., 2019), sin distinguir su comportamiento en distintas estructuras sintácticas. No obstante, para alcanzar una adecuada caracterización del repertorio de colocaciones en L2, estimamos pertinente considerar también la configuración interna de las mismas. Esto es así, principalmente, porque el desarrollo y dominio de las colocaciones en aprendices de L2 está determinado, entre otros factores, por la complejidad sintáctica de las colocaciones (Bardovi-Harlig y Stringer, 2017; Gitsaki, 1996). Además, estudios previos realizados en inglés, francés y holandés como L2 (Paquot, 2018, 2019; Vandeweerd et

al., 2021; Rubin et al., 2021) coinciden en identificar patrones de desarrollo distinto en función de la estructura sintáctica de las colocaciones. Por ejemplo, Vandeweerd et al. (2021) encontraron que la sofisticación de las colocaciones de verbo-sustantivo utilizadas por aprendices de francés L2 aumentó significativamente con el nivel de competencia, mientras que la sofisticación de otras combinaciones analizadas, a saber, las colocaciones de sustantivo-adjetivo y las de adverbio-adjetivo, se mantuvieron similares en los distintos grupos de aprendices. En el presente estudio se analiza la diversidad y la sofisticación de las colocaciones en dos combinaciones distintas, definidas según las categorías gramaticales implicadas y su función sintáctica: 1) las colocaciones verbo-sustantivo(objeto), caracterizadas por formar un sintagma verbal en el que el sustantivo desempeña la función de complemento directo (p. ej.: *dar un paseo*) o de complemento de régimen (p. ej.: *carecer de dinero*), y 2) las colocaciones formadas por un nombre y un adjetivo que lo modifica (p. ej.: *ayuda mutua*). Estos dos tipos de colocaciones son las más frecuentes tanto en los corpus analizados (Cuberos et al., 2024), como en español en general (Koike, 2001).

Por otra parte, estudios previos sobre el desarrollo del repertorio lingüístico en L1 y L2 coinciden en señalar la necesidad de evaluar el dominio lingüístico de hablantes nativos y no nativos en distintos tipos de texto, por cuanto el desarrollo lingüístico se ve condicionado de forma determinante por la situación comunicativa (Aparici, 2010; Cuberos, 2019; Ravid y Tolchinsky, 2002; Rosado et al., 2014). A pesar de ello, la mayoría de los trabajos previos analiza el uso de colocaciones en producciones escritas sin considerar la influencia de la tipología textual. En este estudio, analizamos el uso de colocaciones en dos tipos de textos escritos: narrativos y expositivos. Estos géneros discursivos resultan idóneos para analizar el desarrollo del repertorio lingüístico porque contrastan por sus características textuales definitorias y, además, se dominan en momentos evolutivos distintos (Aparici, 2010).

1.1. *De nuevo sobre el concepto de colocación*

El fenómeno lingüístico al que remite el término *colocación* se ha definido de diversas maneras tanto en lingüística como en el ámbito de la enseñanza de lenguas, y se ha denominado también con términos como unidades léxicas multipalabra o combinaciones prefabricadas, entre otros³. Tanto es así que, como señala Nesselhauf

3. En este artículo, utilizamos los distintos términos empleados por distintos autores, por ejemplo, colocaciones, adoptado, entre otros, por Cowie (1998) o unidad fraseológica, adoptado, entre otros, por Vinogradov (1947), como equivalentes.

(2004), parece que el único aspecto sobre el que existe consenso generalizado es que el fenómeno hace referencia a algún tipo de relación sintagmática entre un grupo de palabras (p. 11).

Tradicionalmente, la noción de *colocación* se aborda en lingüística desde dos perspectivas principales: el enfoque distribucional o basado en la frecuencia, y el enfoque fraseológico (Nesselhauf, 2004, p. 11). La diferencia principal entre estos enfoques radica en las propiedades definitorias que atribuyen a las colocaciones: mientras que el primero describe estas unidades lingüísticas en términos de la frecuencia de coocurrencia de palabras (Firth, 1957, p. 182; Sinclair, 1987, p. 320), el segundo se centra en las posibilidades combinatorias de las palabras que las componen en contextos sintácticos específicos, de modo que describe las colocaciones por medio de criterios estrictamente lingüísticos (Cowie, 1998; Hausmann, 1989).

Desde el enfoque distribucional se sostiene que ciertas palabras tienden a aparecer juntas en determinadas combinaciones y la noción de *colocación* refiere específicamente al contexto en el que aparece una palabra (Laso, 2009, p. 25). A partir de la noción de colocación de Firth (1957, p. 179): «you shall know a word by the company it keeps», lingüistas como Halliday (1961, p. 276) y Sinclair (1991, p. 170) destacan la frecuencia de coocurrencia de ciertas palabras teniendo también en cuenta la distancia entre los elementos que conforman la colocación. De acuerdo con esta aproximación, las colocaciones se identifican aplicando criterios de frecuencia (Sinclair, 1987, p. 320) y también considerando la fuerza de asociación de los elementos que conforman la colocación (Kilgarriff, 2006, p. 38). Por ejemplo, la fuerza de asociación del sintagma *coche oficial* resulta débil porque ambas palabras pueden aparecer en numerosas combinaciones, mientras que, por ejemplo, la colocación *piedra angular* muestra una fuerza de asociación fuerte porque el adjetivo *angular* no suele coaparecer con otras palabras para transmitir un significado unitario.

Los análisis realizados desde esta perspectiva han sido, sin duda, de gran utilidad para comprender mejor las relaciones sintagmáticas que se establecen entre distintos tipos de expresiones multipalabra (Laso, 2009, p. 24). Además, han facilitado la identificación de combinaciones léxicas que desde otros enfoques habrían pasado desapercibidas (Granger y Paquot, 2008, p. 14). A pesar de ello, desde el enfoque fraseológico, autores como Howarth (1996, p. 25) o Mel'čuk (1998, p. 24) sostienen que es fundamental considerar criterios lingüísticos. Así, en el marco del enfoque fraseológico se sostiene que las colocaciones deben necesariamente definirse mediante criterios gramaticales, léxicos y/o semánticos, en lugar de basarse exclusivamente en la frecuencia de coaparación. Desde esta perspectiva se adopta, por tanto, una estrategia

de identificación de colocaciones guiada por criterios lingüísticos previamente establecidos. Así, las unidades léxicas multipalabra se sitúan en un continuo que incluye desde combinaciones libres de palabras semánticamente transparentes en un extremo, hasta expresiones idiomáticas cuyo significado no es composicional en el otro (Cowie, 1998, p. 6). El grado de fijación de las combinaciones se define también como gradual (Corpas, 1996, p. 20), de modo que el *continuum* en el que se ubican las colocaciones implica, sin embargo, áreas difusas cuyos límites resultan difíciles de definir. A pesar de ello, de manera generalizada, se sostiene que las colocaciones son combinaciones prototípicas de dos palabras léxicas que mantienen algún tipo de relación gramatical entre sí (Mittmann, 2013, p. 500) y entre las que se establece un vínculo semántico (Corpas, 2001, p. 42). Si bien los dos miembros de la colocación contribuyen a su significado, uno de ellos (la base) es semánticamente autónomo mientras que el significado del otro (el colocativo) viene semánticamente restringido y seleccionado por la base (Hausmann, 1998, p. 65). Por ejemplo, en la colocación *dar apoyo*, el significado del verbo *dar* (el colocativo) está contextualmente condicionado por el del sustantivo *apoyo* (la base).

Junto a estos dos enfoques tradicionales sobre el concepto de colocación, la propuesta de Bosque (2001, 2011) nos ofrece un marco para analizar las colocaciones desde una perspectiva gramatical en el que defiende que las colocaciones no forman parte del ámbito de la fraseología, sino que se sitúan en la interfaz léxico-sintáctica. Sostiene, además, que la frecuencia por sí misma no proporciona información sobre las propiedades lingüísticas de las colocaciones. En línea con el enfoque fraseológico, Bosque también refiere a la relación semántica que se establece entre los elementos de una colocación; sin embargo, recurre al concepto de 'selección léxica' para defender que es el colocativo, y no la base, el que selecciona palabras de una determinada clase léxica. Siguiendo con el ejemplo anterior, Bosque sostiene que es el verbo *dar* el que selecciona a palabras semánticamente relacionadas como *apoyo*, *ayuda* o *soporte* (Bosque, 2006, p. 382), y no a la inversa.

En este estudio, y en línea con autoras como Granger y Paquot (2008, p. 14) o Laso (2009, p. 9), sostenemos que estos enfoques complementarios son complementarios y no excluyentes, puesto que contribuyen a una comprensión más exhaustiva de las relaciones sintagmáticas que se establecen entre los constituyentes de las unidades léxicas multipalabra. Además, como destaca Wood (2020, p. 42), la combinación de criterios favorece la identificación de unidades fraseológicas. En la mayoría de los estudios basados en corpus, las colocaciones se identifican a partir de criterios distribucionales o fraseológicos y, solo en algunas ocasiones, las investigaciones aplican

ambos criterios (véase, por ejemplo, García-Salido et al., 2019). En este último caso, el proceso suele consistir en la extracción de las colocaciones utilizando medidas estadísticas de frecuencia y, posteriormente, en la desambiguación manual de los resultados para descartar aquellas combinaciones que no cumplan con los criterios fraseológicos previamente establecidos. Adoptar en primer lugar criterios de frecuencia, no obstante, tiene ciertas limitaciones, especialmente cuando se aplica a conjuntos pequeños de datos, puesto que las colocaciones con solo una o dos ocurrencias pueden fácilmente pasar desapercibidas (Wood, 2020). Así pues, en línea con trabajos anteriores (Cuberos, 2019; Cuberos et al., 2024), adoptamos una definición deliberadamente amplia de colocación basada en el enfoque fraseológico que se complementa, por una parte, con un criterio lingüístico de autoridad y, por otra, con un criterio de frecuencia (v. Método). Así, con el objetivo de superar las limitaciones inherentes tanto a los criterios exclusivamente fraseológicos como a los basados únicamente en la frecuencia, el enfoque que adoptamos pretende ofrecer una comprensión más completa del comportamiento de las combinaciones léxicas multipalabra.

2. OBJETIVOS

En este estudio se analiza la diversidad y la sofisticación de las colocaciones en producciones escritas de estudiantes universitarios nativos y no nativos de español. El objetivo es determinar si la diversidad y la sofisticación de estas colocaciones varían según el nivel de competencia de los participantes y el tipo de texto producido. En particular, se examina las colocaciones más frecuentes en español, colocaciones de verbo-sustantivo(objeto) y colocaciones de sustantivo-adjetivo, en dos tipos de textos escritos: narrativos y expositivos.

3. MÉTODO

3.1. *Participantes*

En el estudio participaron 50 estudiantes universitarios: 30 hablantes no nativos de español como L2 (18 mujeres) (rango de edad: 20 – 43;6) y 20 hablantes nativos (14 mujeres) (rango de edad: 18;11 – 28;4). Los participantes de L2 tenían como L1 árabe, chino o coreano. Para determinar su nivel de competencia, docentes experimentados de español como L2 evaluaron y clasificaron los textos en niveles, que fueron distribuidos en dos grupos: inicial-intermedio (A2 y B1) ($n = 11$) e intermedio-avanzado (B2 y C1) ($n = 19$).

3.2. *Tareas de producción*

Los datos proceden de dos corpus análogos: el corpus GRERLI-ES1 (Tolchinsky y Aparici, 2015) y el corpus GRERLI-ES2 (Tolchinsky et al., 2023) los cuales incluyen textos narrativos y expositivos escritos y orales producidos por hablantes nativos y no nativos de español, respectivamente. Las tareas de producción fueron equivalentes para L1 y L2 (v. Berman y Verhoeven, 2002)⁴. Los participantes vieron un vídeo sin diálogos que mostraba situaciones de conflicto en un entorno escolar. A continuación, se les pidió que relataran oralmente y por escrito una experiencia personal similar (texto narrativo) y que discutieran, también oralmente y por escrito, sobre el tema del vídeo (texto expositivo). Como resultado, cada participante produjo cuatro tipos de textos: narrativo escrito, narrativo oral, expositivo escrito y expositivo oral. Para este estudio, se seleccionan únicamente los textos escritos.

3.3. *Identificación de colocaciones*

Las colocaciones se identificaron manualmente mediante la Prueba para la Identificación de Unidades Fraseológicas (TIPUS, por sus siglas en inglés) (Cuberos et al., 2024). TIPUS es un procedimiento para la identificación de colocaciones que combina tres tipos de criterios: fraseológicos, distribucionales y de autoridad lingüística. Para que una unidad léxica se considere una colocación, se deben cumplir dos de los tres criterios establecidos.

El primer criterio consiste en identificar manualmente unidades léxicas candidatas a ser consideradas colocación siguiendo una definición deliberadamente amplia basada en el enfoque fraseológico (véase apartado 1.1.). En concreto, en línea con Corpas (1996; 2001), entendemos las colocaciones como combinaciones sintagmáticas recurrentes entre las que se establece un vínculo léxico-semántico. Para guiar este proceso de identificación, consideramos dos propiedades fundamentales de las colocaciones propuestas por Higueras (2006):

1. Unidades multipalabra compuestas por al menos dos lexemas.
2. Unidades multipalabra con un significado unificado, cierta fijación y coherencia interna.

4. Más detalles sobre el procedimiento de elicitation de los textos puede consultarse en la página web de CHILDES (GRERLI-ES1: <https://childe.talkbank.org/access/Spanish/GRERLI.html>; GRERLIES2: <https://slabank.talkbank.org/access/Spanish/GRERLI-ES2.html>).

Solo se anotaron aquellas combinaciones léxicas canónicas en español, por lo que las colocaciones creativas y/o erróneas de los aprendices (p. ej., *tener un paseo*^{*}, por *dar un paseo*), no se identificaron como candidatas a colocación, principalmente, porque no cumplen con los criterios fraseológicos establecidos, pero también porque estas colocaciones que se desvían de la norma no forman parte del objeto de estudio. Para garantizar la fiabilidad de los *juicios intuitivos* (Wray, 2008) que guían este proceso, se discutieron los casos dudosos entre las investigadoras y se calculó la fiabilidad de la anotación en el 20% de los textos mediante el coeficiente Alfa de Cronbach donde se obtuvo un acuerdo significativo entre juezas ($\alpha \geq .784$) (Field, 2009). Como resultado de este primer paso, se anotaron las combinaciones léxicas candidatas a colocación. A modo de ejemplo, incluimos a continuación algunos ejemplos para ilustrar las combinaciones léxicas anotadas en este primer paso:

- (1) *Si lo ha hecho mal, que pida perdón*
- (2) *Hay tantas imágenes violentas como en la guerra o en conflictos laborales*
- (3) *Gente no le gusta sentarse al lado porque tiene envidia*
- (4) *Nos cogían como boxeadores y nos daban masajes pequeñitos en la espalda*

El segundo paso, basado en el enfoque distribucional, consiste en medir la fuerza de asociación y frecuencia de coocurrencia de las unidades léxicas que conforman las combinaciones candidatas a colocación mediante la puntuación *logDice*. Se trata de una medida de asociación léxica que combina información sobre la frecuencia de coocurrencia y la fuerza de asociación de los constituyentes de la colocación (Rychlý, 2008). Estas puntuaciones se obtuvieron mediante la aplicación Word Sketch disponible en el programa de análisis de corpus Sketch Engine (Kilgariff et al., 2014). Las búsquedas, que se realizan tomando como referencia la base de la colocación, ofrecen listados organizados según las relaciones gramaticales que se establecen entre la base y los colocativos. Por ejemplo, en la búsqueda de *línea*, sus colocativos aparecen agrupados según se trate de verbos que coocurren con la base como objeto directo (p. ej., *trazar; trazar una línea*), o de adjetivos que la modifican (p. ej., *recto; línea recta*). Decidimos utilizar la medida de asociación *logDice* en lugar de otras alternativas como la puntuación *MI* o la puntuación *T* porque su valor no depende del tamaño del corpus y porque, a diferencia de estas otras medidas, se basa en un máximo teórico, lo que resulta fundamental para establecer comparaciones entre distintos corpus (Gablasova et al., 2017). Se ha demostrado, además, que *logDice* funciona mejor que otras medidas en la identificación de colocaciones (Gablasova et al., 2017). En concreto, a las combinaciones candidatas

se les asignó la puntuación *logDice* obtenida en el corpus de español esTenTen (Kilgariff y Renau, 2013), disponible también en Sketch Engine. El máximo teórico de esta puntuación es 14, sin embargo, los valores que alcanza suelen ser inferiores a 10 (Rychlý, 2008). Por esta razón convinimos que las combinaciones candidatas a colocación debían obtener una puntuación *logDice* igual o superior a 5 para cumplir con el criterio distribucional. En los casos en que colocaciones identificadas en el primer paso son instancias de la misma colocación (p. ej., *sacar buenas notas* y *sacar malas notas*), la puntuación se calculó a partir de la media de las puntuaciones de cada combinación. La Tabla 1 recoge las colocaciones identificadas en el primer paso junto con sus puntuaciones *logDice*, aquellas sombreadas en gris no alcanzaron el umbral mínimo de 5 puntos.

Colocaciones candidatas	Puntuación <i>logDice</i>
Pedir perdón	9.97
Conflictivo laboral	7.63
Dar un masaje	4.23
Tener envidia	2.99

Tabla 1. Puntuaciones *logDice* de las colocaciones candidatas

Por último, como criterio lingüístico de autoridad, se verifica la presencia de las colocaciones en el *Diccionario combinatorio práctico del español contemporáneo (Práctico)* (Bosque, 2006). Decidimos utilizar este diccionario porque está basado en la noción de colocación de Bosque (2011) que, recordamos, sostiene que las colocaciones responden a casos de selección léxica en las que el collocativo, interpretado en sus obras como predicado, selecciona la base, interpretada en sus obras como argumentos. De este modo, las búsquedas en *Práctico* se realizaron tomando como referencia el collocativo en lugar de la base.

Una vez asignadas las puntuaciones *logDice* a las colocaciones candidatas y tras verificar su presencia en *Práctico*, se identificó la lista final de colocaciones. Si bien encontramos casos en los que algunas colocaciones no estaban recogidas en *Práctico* (p. ej., *conflicto laboral*) o no alcanzaban el umbral de frecuencia establecido (p. ej., *tener envidia*), solo se descartaron aquellas colocaciones que no cumplían ninguno de los dos criterios.

Como resultado de aplicar este procedimiento, en los corpus de análisis se identificaron 1392 colocaciones: 710 en el corpus de L1 y 682 en el corpus de L2. En este estudio, como se ha dicho, se analizan las colocaciones de

verbo-sustantivo(objeto) (p. ej., *ocasionar problemas*; *ir al teatro*) y las de sustantivo-adjetivo (p. ej., *año pasado*; *ayuda mutua*), que son los dos tipos de colocaciones más frecuentes en español y en los corpus analizados. En la Tabla 2 se muestra el número total de colocaciones y el número de colocaciones de cada tipo por nivel de competencia.

Nivel de competencia	N.º verbo-sustantivo	N.º sustantivo-adjetivo	N.º total
Inicial-intermedio	72	31	167
Intermedio-avanzado	189	77	515
Nativo	321	90	710

Tabla 2. Número de colocaciones total y número de verbo-sustantivo(objeto) y sustantivo-adjetivo por nivel de competencia

3.3. Complejidad fraseológica

La complejidad fraseológica es, como se ha dicho, un constructo que describe el uso de dos propiedades del uso que se hace de las colocaciones: la diversidad colocacional y la sofisticación colocacional (Paquot, 2019). La diversidad, que refiere a la variedad de colocaciones utilizadas, se obtuvo mediante la proporción de colocaciones diferentes (lemas)⁵ sobre la raíz cuadrada del número total de colocaciones (Paquot, 2019). Por su parte, la sofisticación, que refiere al uso de colocaciones menos comunes y avanzadas, se calculó mediante las puntuaciones *logDice*. Esta medida de asociación léxica combina información sobre la frecuencia de coaparación y la fuerza de asociación de los elementos que conforman la colocación (Rychlý, 2008), de tal modo que aporta información sobre cuán frecuente es la combinación al mismo tiempo que indica cuál es la fuerza de la asociación entre ambas palabras (Gablasová et al., 2017). El valor medio que suele alcanzar esta puntuación es 10 (Rychlý, 2008), por lo que los valores por encima de 7 se consideran indicativos de colocaciones fuertes (Cuberos et al., 2024). Decidimos calcular la sofisticación de las colocaciones mediante una medida de fuerza de asociación porque la gran mayoría de colocaciones que obtienen puntuaciones altas en estas medidas tienen significados específicos (Ellis et al., 2008;

5. En línea con García-Salido y García (2018) y Cuberos et al. (2024), optamos por calcular la diversidad de las colocaciones considerando los lemas en lugar de las formas, como es habitual en lenguas como el inglés. Esta decisión se justifica tanto por la complejidad morfológica del español, que puede inflar artificialmente la diversidad al contar formas flexionadas, como por los hallazgos en procesamiento de colocaciones que indican que la variación morfológica no inhibe su ventaja de procesamiento (Vilkaitė-Lozdienė, 2022; Pérez-Serrano e Hidalgo-Gallardo, en prensa).

Gablasova et al., 2017). Por ejemplo, mientras que *perro grande* y *perro pequeño* son las combinaciones de sustantivo-adjetivo más frecuentes en el corpus esTenTen23 (Kilgariff y Renau, 2013), las combinaciones *perro guardián* y *perro callejero* son las que obtienen puntuaciones logDice más altas, lo que indica que las colocaciones con puntuaciones más altas tienen una mayor especialización semántica. Además, porque en estudios anteriores (véase, por ejemplo, Paquot, 2019; Vandeweerd et al., 2021; Rubin et al., 2021) estas medidas han resultado especialmente útiles para describir el desarrollo de la competencia fraseológica en L2.

3.4. Análisis de datos

Para abordar los objetivos de investigación, en primer lugar, se realizaron análisis estadísticos descriptivos y exploraciones visuales de los datos con diagramas de cajas. Seguidamente, se realizaron una serie de análisis ANOVA robustos de medidas repetidas con el Nivel de competencia como factor intersujeto (intermedio, avanzado, nativo) y el Tipo de texto como factor intrasujeto (narrativo, expositivo). Los análisis se llevaron a cabo en R (R Core Team, 2021) con la función *bwtrim* del paquete *WSR2* (Mair y Wilcox, 2016). En caso de interacciones significativas, se analizaron efectos simples de Nivel de competencia y de Tipo de texto con las funciones *lincon* y *yuend*, respectivamente. En las comparaciones por pares se aplicó la corrección de Bonferroni para evitar errores de tipo I.

4. RESULTADOS

En este apartado se recoge, en primer lugar, los resultados relativos a la complejidad fraseológica de las colocaciones verbo-sustantivo(objeto). Seguidamente, se incluyen los resultados relativos a la complejidad fraseológica de las colocaciones sustantivo-adjetivo.

4.1. Complejidad fraseológica de las colocaciones verbo-sustantivo(objeto) en textos narrativos y expositivos por nivel de competencia

En la Tabla 3 se muestra la estadística descriptiva de las medidas de complejidad fraseológica analizadas de las colocaciones verbo-sustantivo(objeto) en los textos narrativos y expositivos en L1 y L2. En la Figura 1 se ilustran gráficamente los índices de diversidad y sofisticación por nivel de competencia en ambos tipos de texto. Los resultados descriptivos indican, por una parte, que tanto la diversidad como la sofisticación de estas colocaciones son similares en los dos tipos de textos

en todos los niveles de competencia y, por otra, que ambas aumentan con el nivel de competencia. A juzgar por las desviaciones típicas de las medidas de complejidad fraseológica, la diversidad de las colocaciones presenta una baja variabilidad en todos los grupos analizados, aunque es algo superior en los hablantes de nivel intermedio. En contraste, la sofisticación presenta una alta variabilidad, sobre todo en los grupos de hablantes no nativos.

Medida	Nivel	M (DT)	Narrativo M (DT)	Expositivo M (DT)
Diversidad	Intermedio	1,09 (0,69)	1,17 (0,65)	1,02 (1,05)
	Avanzado	1,55 (0,62)	1,58 (0,69)	1,52 (0,56)
	Nativo	2,54 (0,60)	2,53 (0,74)	2,55 (0,47)
Sofisticación	Intermedio	5,96 (4,38)	7,03 (3,74)	4,89 (4,88)
	Avanzado	7,86 (2,38)	7,81 (2,47)	7,9 (2,35)
	Nativo	8,26 (1,17)	7,91 (1,36)	8,61 (0,88)

Tabla 3. Estadística descriptiva (Media (M) y Desviación típica (DT)) de la diversidad y la sofisticación de las colocaciones verbo-sustantivo en textos narrativos y expositivos por nivel de competencia

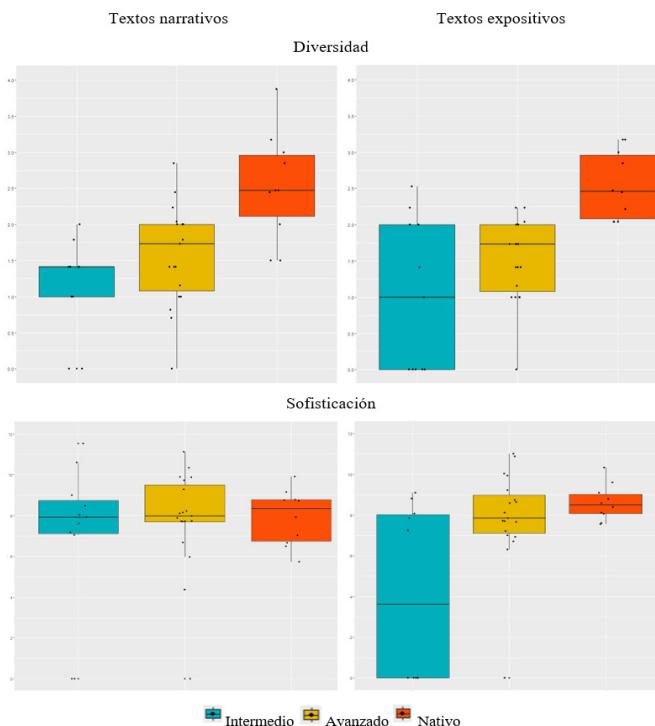


Figura 1. Distribución de la diversidad y la sofisticación de las colocaciones verbo-sustantivo en textos narrativos y expositivos por nivel de competencia

Los resultados del ANOVA factorial para la diversidad indican que se da un efecto simple del Nivel de competencia, $Q_{(2,13,36)} = 19,77, p < ,001$, en el sentido de que a medida que aumenta el nivel de competencia, aumenta la diversidad léxica. Los análisis *post-hoc* indican que se dan diferencias en todas las comparaciones por pares de modo que la diversidad de las colocaciones verbo-sustantivo aumenta progresivamente con el nivel de competencia (en todas las comparaciones $p < ,016$). En el caso de las diferencias en función del tipo de texto, no encontramos un efecto simple del Tipo de texto ni tampoco una interacción significativa entre este factor y el Nivel de competencia, de esta forma los resultados indican que la diversidad de este tipo de colocaciones es comparable en textos narrativos y expositivos.

Por su parte, los resultados del ANOVA para la sofisticación de las colocaciones indican que en este caso también se da un efecto simple del Nivel de competencia, $Q_{(2,74)} = 4,14, p = ,011$, por el cual aumenta la sofisticación léxica de las colocaciones con el nivel de competencia. En este caso, sin embargo, los análisis *post-hoc* indican que se dan diferencias en la comparación del grupo de nivel intermedio con el grupo de nivel avanzado ($p = ,013$) y con el grupo de nativos ($p = ,011$), en el sentido de que los hablantes más expertos incorporan en sus textos colocaciones más sofisticadas, mientras que las diferencias entre los hablantes de nivel avanzado y los nativos no resultan significativas. Tampoco en este caso se da un efecto simple del Tipo de texto ni se observa una interacción significativa entre ambos factores. Esto supone que, de nuevo, tampoco hay diferencias en la sofisticación de estas colocaciones en los textos narrativos y expositivos.

4.2. Complejidad fraseológica de las colocaciones sustantivo-adjetivo en textos narrativos y expositivos por nivel de competencia

La Tabla 4 recoge la estadística descriptiva de las medidas de complejidad fraseológica analizadas de las colocaciones de sustantivo-adjetivo. En la Figura 2 se ilustran gráficamente los índices de diversidad y sofisticación por nivel de competencia en ambos tipos de texto. Al igual que en las colocaciones verbo-sustantivo, la complejidad fraseológica de estas colocaciones tampoco parece depender del tipo de texto, pues los valores de diversidad y sofisticación léxicas son similares en textos narrativos y expositivos. Se observan, eso sí, patrones de desarrollo distintos por nivel de competencia para cada una de las dimensiones analizadas. Así, al igual que en las colocaciones de verbo-sustantivo, la diversidad de estas colocaciones también aumenta en función del nivel de competencia. En cuanto a la sofisticación, los resultados descriptivos muestran diferencias entre hablantes nativos

y no nativos; sin embargo, los índices medios de sofisticación son similares entre los hablantes no nativos. En ambos casos, las desviaciones típicas son más altas que las observadas en las colocaciones de verbo-sustantivo. Esta variabilidad se acentúa sobre todo en la sofisticación de las colocaciones de los hablantes no nativos. Por el contrario, los hablantes nativos muestran un uso más homogéneo, pues los valores de la desviación típica son más bajos que los observados en los otros grupos.

Medida	Nivel	M (DT)	Narrativo M (DT)	Expositivo M (DT)
Diversidad	Intermedio	0,69 (0,63)	0,56 (0,56)	0,83 (0,69)
	Avanzado	0,91 (0,74)	0,83 (0,81)	1 (0,68)
	Nativo	1,53 (0,55)	1,24 (0,55)	1,81 (0,39)
Sofisticación	Intermedio	4,23 (3,83)	3,71 (3,9)	4,75 (3,87)
	Avanzado	4,97 (3,87)	3,95 (3,63)	5,99 (3,93)
	Nativo	6,85 (2,36)	6,62 (3,24)	7,08 (1,09)

Tabla 4. *Estadística descriptiva (Media (M) y Desviación típica (DT) de la diversidad y la sofisticación de las colocaciones verbo-sustantivo en textos narrativos y expositivos por nivel de competencia*

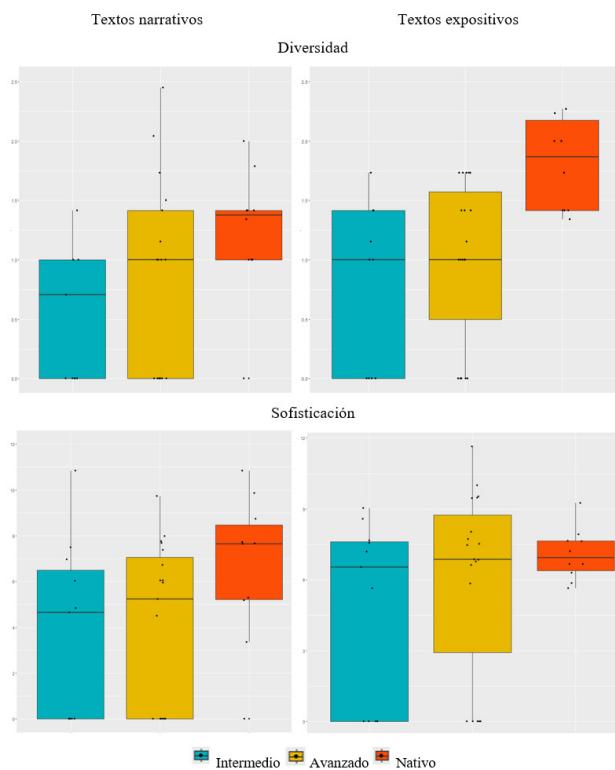


Figura 2. Distribución de la diversidad y la sofisticación de las colocaciones sustantivo-adjetivo en textos narrativos y expositivos por nivel de competencia

Los resultados del ANOVA factorial de la diversidad de las colocaciones de sustantivo-adjetivo indican un efecto simple del Nivel de competencia, $Q_{(2, 12, 94)} = 8,36$, $p = ,004$, en el sentido de que a medida que aumenta el nivel de competencia aumenta también la variedad de colocaciones utilizadas. En este caso, y a diferencia de lo observado en las colocaciones de verbo-sustantivo, los análisis *post-hoc* indican que solo hay diferencias significativas en las comparaciones entre los hablantes nativos y los grupos de nivel intermedio ($p = ,003$) y avanzado ($p = ,007$), mientras que las diferencias entre los dos grupos de no nativos no resultan significativas. Al igual que en las colocaciones de verbo-sustantivo, tampoco se da un efecto simple del Tipo de texto ni una interacción significativa entre ambos factores. De este modo, la diversidad de las colocaciones de sustantivo-adjetivo también se mantiene comparable en los textos narrativos y expositivos.

En cambio, y a diferencia de lo observado para las colocaciones de verbo-sustantivo, los resultados del ANOVA factorial de la sofisticación de estas colocaciones de sustantivo-adjetivo indican que no hay efectos simples del Nivel de competencia ni del Tipo de texto. Tampoco se ha encontrado una interacción significativa entre ambos factores. Estos resultados indican, por tanto, que la sofisticación de las colocaciones sustantivo-adjetivo es similar en todos los niveles de competencia analizados y que se mantiene comparable en los textos narrativos y expositivos.

5. DISCUSIÓN

En este estudio hemos analizado dos propiedades fundamentales de la complejidad fraseológica: la diversidad y la sofisticación léxica de las colocaciones en producciones escritas de estudiantes universitarios que son hablantes nativos y no nativos de español y que, en el caso de estos últimos, presentan diferentes niveles de competencia. Estas propiedades se han examinado en los dos tipos de colocaciones más frecuentes en español: verbo-sustantivo(objeto) y sustantivo-adjetivo, en dos tipos de textos escritos: narrativos y expositivos.

Según nuestros resultados, y en línea con lo observado en estudios previos, la complejidad fraseológica aumenta con el grado de dominio lingüístico, por cuanto los hablantes más experimentados producen una mayor diversidad de colocaciones, las cuales son, además, más sofisticadas. Este desarrollo se da con independencia del propósito comunicativo, pues no encontramos diferencias significativas en función del tipo de texto, y el índice de diversidad y sofisticación de las colocaciones utilizadas se mantienen en los textos narrativos y expositivos en todos los grupos

de hablantes estudiados. Estos resultados proporcionan evidencia adicional de que el uso de colocaciones constituye un indicador fundamental de competencia lingüística en L2 (Cuberos, 2019; Cuberos et al., 2024; Fan, 2009; García-Salido et al., 2019), y esto es así con independencia del género discursivo (Cuberos, 2019). Los patrones de desarrollo identificados sí varían según la estructura sintáctica de las colocaciones. Este resultado, atestiguado en estudios previos (Paquot, 2018; 2019; Vandeweerd et al., 2021), subraya la importancia de considerar la naturaleza sintáctica de las colocaciones para analizar apropiadamente el desarrollo de la competencia fraseológica en L2. A continuación, discutimos en primer lugar los resultados relativos a las colocaciones verbo-sustantivo y, seguidamente, hacemos lo propio con las de sustantivo-adjetivo.

En relación con las colocaciones de verbo-sustantivo, los resultados indican que la complejidad fraseológica aumenta con el nivel de competencia si bien se observan patrones de desarrollo distintos para cada una de las dimensiones analizadas. Así, la diversidad de las colocaciones utilizadas muestra un aumento progresivo con diferencias significativas en todas las comparaciones por pares mientras que, en el caso de la sofisticación, se observan diferencias significativas en la comparación entre los hablantes no nativos de niveles intermedios con los de nivel avanzado y los nativos, entre los cuales no hay diferencias. Un primer análisis pormenorizado de estos resultados nos lleva a hipotetizar que algunas de estas colocaciones incluyen lo que la literatura previa (véase, por ejemplo, García Salido, 2014; 2016; Pérez-Serrano y Mavrou, 2022) denomina construcciones con verbo de apoyo o verbo ligero, como *hacer fotos*, *dar ejemplo* o *tener confianza*, es decir, colocaciones en las que el verbo está deslexicalizado y el contenido semántico lo aporta principalmente el sustantivo. Ejemplos de este tipo de construcciones son *hacer fotos*, *dar ejemplo* o *tener confianza*. Estudios previos señalan que los aprendices de L2 tienden a utilizar estas construcciones con verbo de apoyo con más frecuencia que los nativos; estos, por su parte, disponen de un repertorio más variado de este tipo de formas (García Salido, 2014). Asimismo, en un trabajo anterior en el mismo corpus analizado en este estudio (Cuberos et al., 2024), observamos que hablantes nativos y no nativos de distintos niveles de competencia utilizan frecuentemente estas construcciones, pero que los aprendices más experimentados incorporan de forma recurrente otras combinaciones con verbos con significado pleno (p. ejemplo, *acarrear consecuencias*). Dado el papel central que parecen desempeñar estas construcciones en el desarrollo de la competencia fraseológica en L2, estas merecen, sin duda, un análisis pormenorizado en próximos estudios.

En el caso de las colocaciones sustantivo-adjetivo encontramos un escenario ligeramente distinto. Si bien los nativos utilizan un repertorio más variado de este tipo de colocaciones que los no nativos de distintos niveles de competencia, entre los que no se dan diferencias significativas, la sofisticación de estas construcciones es similar en todos los grupos analizados, pues no hay un efecto de nivel de lengua. Estos resultados sugieren que, a diferencia de lo que ocurre con las colocaciones de verbo-sustantivo, el nivel de competencia parece jugar aquí un papel menos relevante. Este comportamiento diferenciado según la estructura sintáctica, y atestiguado también en otras lenguas como inglés (Paquot, 2019) y francés (Vandeweerd et al., 2021), puede deberse, por una parte, a que se trata de colocaciones menos frecuentes en español (Cuberos et al., 2024; Koike, 2001), lo que podría condicionar su desarrollo incluso en niveles avanzados. Por otra parte, la propia naturaleza sintáctica de estas colocaciones podría también ser responsable de la ausencia de diferencias significativas entre los grupos.

En las colocaciones sustantivo-adjetivo, la coaparación de ambos constituyentes parece tener, en algunos casos, un nivel de fijación menor de tal manera que presentan, de algún modo, mayor opcionalidad que las combinaciones verbo-sustantivo, especialmente las de verbo de apoyo. Así, mientras el sustantivo y el verbo se encuentran aquí en una relación de interdependencia en virtud de la cual, por así decirlo, el sustantivo depende gramaticalmente del verbo y el verbo necesita semánticamente al sustantivo, en las colocaciones sustantivo-adjetivo, este último aporta, por definición, información adicional. Este carácter opcional otorgaría a los adjetivos una mayor flexibilidad en su uso. En futuros trabajos sería, sin duda, conveniente profundizar en la naturaleza semántica y el comportamiento discursivo de este tipo de colocaciones.

Tomados en conjunto, los resultados de este estudio subrayan, por una parte, la naturaleza multifacética de las colocaciones y apuntan a que las propiedades definitorias de las mismas interactúan entre sí para describir el desempeño lingüístico en L2 (Paquot, 2019). Además, confirman la idoneidad de considerar la diversidad de las colocaciones utilizadas en un texto para analizar y caracterizar el desarrollo lingüístico en L2 (Cuberos et al., 2024) y apuntan, como novedad, a la relevancia que supone examinar la estructura sintáctica de estos elementos. Concretamente, el análisis de las colocaciones verbo-sustantivo resulta especialmente relevante para capturar diferencias en el desarrollo de la competencia fraseológica como parte de la competencia léxica de los aprendices. Por último, nos parece importante señalar cómo, en futuros trabajos, sería pertinente analizar su comportamiento en

relación con otros indicadores de desarrollo lingüístico, tales como la complejidad sintáctica o la densidad discursiva (Paquot, 2019). Igualmente, se podrían estudiar otras formas de medir la complejidad fraseológica, por ejemplo, explorando otros índices de sofisticación.

DECLARACIÓN DE CONTRIBUCIÓN DE AUTORÍA

Las autoras han contribuido de manera equitativa y significativa a la elaboración del artículo. Las tareas principales realizadas por cada autora se listan a continuación:

Rocío Cuberos Vicente: Conceptualización, Investigación, Metodología, Diseño de instrumento de identificación de colocaciones, Validación del instrumento de identificación de colocaciones, Análisis formal de datos, Redacción–borrador original, Redacción–revisión y edición.

Elisa Rosado Villegas: Conceptualización, Investigación, Metodología, Diseño de instrumento de identificación de colocaciones, Validación del instrumento de identificación de colocaciones, Supervisión, Validación, Redacción–revisión y edición.

Ambas autoras han participado en la revisión crítica del documento, aprobaron la versión final del manuscrito y asumen responsabilidad por su contenido.

FINANCIACIÓN

La recogida de los datos utilizados en esta investigación ha sido financiada por el proyecto «El desarrollo de la calidad textual en castellano y catalán L2. Indicadores léxicos, morfosintácticos y discursivos» (I.P.: J. Perera, Universitat de Barcelona) (EDU2012-34394, Ministerio de Economía y Competitividad). El tratamiento y análisis de los datos ha sido financiado por la beca postdoctoral Margarita Salas otorgada a R. Cuberos por la Universitat de Barcelona (Next Generation EU y el Ministerio de Universidades).

REFERENCIAS

- Aparici, M. (2010). *El desarrollo de la conectividad discursiva en diferentes géneros y modalidades de producción* [Tesis doctoral inédita, Universitat de Barcelona].
- Bardovi-Harlig, K., y Stringer, D. (2017). Unconventional expressions: Productive syntax in the L2 acquisition of formulaic language. *Second Language Research*, 33(1), 61-90. doi: <https://doi.org/10.1177/02676583166417>
- Berman, R. y Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text production abilities: Speech and writing. *Written Language and Literacy*, 5(2), 1-43. doi: <https://doi.org/10.1075/wll.5.1.02ber>
- Bosque, I. (2001). Sobre el concepto de ‘colocación’ y sus límites. *Lingüística Española Actual*, 23(1), 9-40.
- Bosque, I. (2006). *Diccionario combinatorio práctico del español contemporáneo*. SM.

- Bosque, I. (2011). Deducing collocations. En I. Boguslavsky y L. Wanner (Eds.), *Proceedings of the 5th International Conference on Meaning-Text Theory* (pp. 6-13).
- Bosque, I. (2017). On the conceptual bases of collocations: Restricted adverbs and lexical selection. En S. Torner y E. Bernal (Eds.), *Collocations and other lexical combinations in Spanish: Theoretical, lexicographical and applied perspectives* (pp. 9-20). Routledge.
- Boone, G., y Eyckmans, J. (2023). Exploring collocation development in L2 German from students' perspective: A contrasting case study. *Studies in Second Language Learning and Teaching*, 13(3), 571-599. doi: <https://doi.org/10.14746/ssllt.32539>
- Corpas, G. (1996). *Manual de fraseología española*. Gredos.
- Corpas, G. (2001). Apuntes para el estudio de la colocación. *Lingüística Española Actual*, 23(1), 41-56.
- Cowie, A. (1998). *Phraseology: Theory, analysis and applications*. Oxford University Press.
- Crossley, S., Cai, Z. y McNamara (2012). Syntagmatic, paradigmatic, and automatic N-gram approaches to assessing essay quality. En *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25*, 214-219.
- Cuberos, R. (2019). *Indicadores léxicos de calidad textual en español nativo y no nativo* [Tesis doctoral, Universitat de Barcelona]. doi: <https://deposit.ub.edu/dspace/handle/2445/178686>
- Cuberos, R., Rosado, E. y Mañas, I. (2024). Towards a graded lexical inventory of multi-word combinations. The distribution of collocations across Spanish CEFR-levels. *International Journal of Applied Linguistics*, 175(1), 46-76. doi: <https://doi.org/10.1075/itl.22021.cub>
- Durrant, P. y Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, 47, 157-177. doi: <https://doi.org/10.1515/iral.2009.007>
- Ellis, N.C., Simpson-Vlach, R., Romer, U., Brook O'Donnell, M. y Wulff, S. (2015). Learner corpora and formulaic language in second language acquisition. In S. Granger, G. Gilquin y F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 357-378). Cambridge University Press. doi: <https://doi.org/10.1017/CBO9781139649414.016>
- Fan, M. (2009). An exploratory study of collocational use by ESL students. A task based approach. *System*, 37, 110-123. doi: <https://doi.org/10.1016/j.system.2008.06.004>
- Firth, J. (1957). *Papers in Linguistics 1934-1951*. Oxford University Press.
- Gablasova, D., Brezina, V. y McEnery, T. (2017). Exploring learner language through corpora: Comparing and interpreting corpus frequency information. *Language Learning*, 67(21), 130-154.
- García-Salido, M. (2016). Error analysis of support verb constructions in written Spanish learner corpora. *The Modern Language Journal*, 100, 362-376. doi: <https://doi.org/10.1111/modl.12320>
- García-Salido, M., García, M. y Alonso-Ramos, M. (2019). Towards a graded dictionary of Spanish collocations. En I. Kosem, T. Zingano, M. Correira, J.P. Ferreira, M.

- Jansen, I. Pereira, J. Kallas, M. Jakubíček, S. Krek y C. Tiberius (Eds.), *Proceedings of eLex 2019 conference* (pp. 849–864).
- García-Salido, M., Lorente, P. y Basanta, A. (2019). Las construcciones con verbos de apoyo del español en la producción escrita de aprendices francófonos. *Journal of Spanish Language Teaching*, 6(1), 32–48.
doi: <https://doi.org/10.1080/23247797.2019.1590022>
- Gitsaki, C. (1996). *The development of ESL collocational knowledge* [Doctoral dissertation, The University of Queensland]. doi: <https://doi.org/10.14264/205374>
- Granger, S. y Paquot, M. (2008). Disentangling the phraseological web. En S. Granger, y F. Meunier (Ed.), *Phraseology: An interdisciplinary perspective* (pp. 28–49). John Benjamins. doi: <https://doi.org/10.1075/z.139.07gra>
- Halliday, M. A. (1961). *Categories of the theory of grammar. Department of English Language and General Linguistic Monographs* (pp. 241–292). Edinburgh University Press.
- Hausmann, F. (1979). Un dictionnaire des collocations est-il possible? *Travaux de Linguistique et de Littérature*, 17(1), 187–195
- Hausmann, F. (1998). O diccionario de colocacións: criterios de organización. In X. Ferro Ruibal (Ed.), *Actas do I Coloquio Galego de Fraseoloxía* (pp. 63–81). Centro Ramón Piñeiro, Xunta de Galicia.
- Higueras, M. (2006). *Las colocaciones y su enseñanza en la clase de ELE*. Arco Libros.
- Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Max Niemeyer Verlag.
doi: <https://doi.org/10.1515/9783110937923>
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4–21. doi: <https://doi.org/10.1016/j.esp.2007.06.001>
- Kecskes, I. (2007). Formulaic language in English Lingua Franca. En I. Kecskes y L. R. Horn (Eds.), *Explorations in Pragmatics* (pp. 191–218). Mouton de Gruyter.
doi: <https://doi.org/10.1515/9783110198843.3.191>
- Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*, 1(2), 263–276. doi: <https://doi.org/10.1515/cllt.2005.1.2.263>
- Kilgarriff, A. y Renau, I. (2013). esTenTen, a vast web corpus of Peninsular and American Spanish. *Procedia-Social and Behavioral Sciences*, 95, 12–19.
doi: <https://doi.org/10.1016/j.sbspro.2013.10.617>
- Kilgarriff, A., Baisa, V., Busta, J., Jakubicek, M., Kovar, V., Michelfet, J., Rychlý, P. y Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography ASIALEX*, 1, 7–36.
- Koike, K. (2001). *Colocaciones léxicas en el español actual: Estudio formal y léxicosemántico*. Universitat de Alicante y Takushoku University.
- Laso, N. (2002). *A corpus-based study of phraseological behaviour of abstract nouns in medical English: a needs analysis of a Spanish medical community*. [Tesis doctoral, Universitat de Barcelona]. doi: <http://hdl.handle.net/10803/1671>
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd Ed.). Lawrence Erlbaum.
- Mair, P. y Wilcox, R. (2016). *WRS2: A collection of robust statistical methods*. doi: <https://CRAN.R-project.org/package=WRS2>

- Mel'čuk, I. (2012). Phraseology in the language, in the dictionary, and in the computer. *Yearbook of Phraseology*, 3, 31–56. doi: <https://doi.org/10.1515/phras-2012-0003>
- Mittmann, B. (2013). New tendencies in the treatment of collocations. En R. H. Gouws, U. Heid, W. Schweickard y H. E. Wiegand (Eds.), *Supplementary volume dictionaries. An International Encyclopedia of Lexicography* (pp. 500–509). De Gruyter Mouton. doi: <https://doi.org/10.1515/9783110238136.500>
- Nesselhauf, N. (2004). What are collocations? En D. Allerton, N. Nesselhauf y P. Skandera (Eds.), *Phraseological units: Basic concepts and their application* (pp. 1–21). Schwabe.
- Paquot, M. (2018). Phraseological competence: A missing component in university entrance language tests? Insights from a study of EFL learners' use of statistical collocations. *Language Assessment Quarterly*, 15(1), 29–43. doi: <https://doi.org/10.1080/15434303.2017.1405421>
- Paquot, M. (2019). The phraseological dimension in interlanguage complexity research. *Second Language Research*, 35(1), 121–145. doi: <https://doi.org/10.1177/0267658317694221>
- Pérez-Serrano, M. y Mavrou, I. (2022). La incidencia del tipo de ejercicio y de la capacidad de la memoria operativa en el aprendizaje de las construcciones con verbo de apoyo en español como segunda lengua. *Onomázein*, 58, 47–68. doi: <https://doi.org/10.7764/onomazein.58.03>
- Pérez-Serrano, M. y Hidalgo-Gallardo, M. (en prensa). ¿Son *da un paseo* y *daba un paseo* variantes de la misma colocación o colocaciones distintas? Variación flexiva y procesamiento de las colocaciones En J. Amenós Pons, C. Chacón García y M. Pérez Serrano (Eds.) *Léxico y sintaxis en la adquisición de español como segunda lengua: perspectivas actuales*. Ediciones Complutense.
- Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing and use*. John Benjamins.
- Sinclair, J. (1987). *Looking up – An account of the COBUILD Project in lexical computing*. Harper-Collins.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- R Core Team (2021). *R: A language environment for statistical computing*. Version 4.1.2.
- Ravid, D. y Tolchinsky, L. (2002). Developing linguistic literacy: a comprehensive model. *Journal of Child Language*, 29(2), 417–447. doi: <https://doi.org/10.1017/S0305000902005111>
- Real Academia Española. (2014). *Diccionario de la lengua española* (23rd Ed.).
- Rosado, E., Salas, N., Aparici, M. y Tolchinsky, L. (2014). Production of linguistic devices for attaining a detached stance in Spanish and Catalan. *Journal of Pragmatics*, 60, 36–53. doi: <https://doi.org/10.1016/j.pragma.2013.10.004>
- Rubin, R., Housen, A. y Paquot, M. (2021). Phraseological complexity as an index of L2 Dutch writing proficiency: A partial replication study. En Granger, S. (Ed.), *Perspectives on the second language phrasicon: The view from learner corpora* (pp. 101–125). Multilingual Matters. doi: <https://doi.org/10.21832/9781788924863-006>

- Rychlý, P. (2008). A lexicographer-friendly association score. En P. Sojka y A. Horák (Eds.) *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN 2008* (pp. 6 - 9). University de Masaryk.
- Tolchinsky, L. y Aparici, M. (2015). *Spanish GRERLI-ES1 corpus* [Conjunto de datos]. CHILDES Talkbank. doi: <https://doi.org/10.21415/T53G63>
- Tolchinsky, L., Rosado, E., Aparici, M. y Cubero, R. (2023). *Spanish GRERLI-ES2 Corpus* [Conjunto de datos]. CHILDES Talkbank.
doi: <https://doi.org/10.21415/2DJF-JH34>
- Vandeweerd, N., Housen, A. y Paquot, M. (2021). Applying phraseological complexity measures to L2 French: A partial replication study. *International Journal of Learner Corpus Research*, 7(2), 197-229.
- Vilkaitė-Lozdienė, L. (2022). Do different morphological forms of collocations show comparable processing facilitation? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 48(9): 1328-1347. doi: <https://doi.org/10.1037/xlm0001130>
- Wood, D. (2020). Classifying and identifying formulaic language. En S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 30-45).
<https://doi.org/10.4324/9780429291586>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
doi: <https://doi.org/10.1017/CBO9780511519772>

RESEÑAS



RESEÑAS

ÁLVAREZ GARCÍA, Esther (ed.), *Dos lenguas, un cerebro: el procesamiento lingüístico en hablantes bilingües*. Berlín: Peter Lang, 2024, 301 pp., ISBN: 978-3-631-89416-3.

El monográfico *Dos lenguas, un cerebro: el procesamiento lingüístico en hablantes bilingües*, editado por Esther Álvarez García, supone un acercamiento sistemático y riguroso al complejo (y fascinante) mundo del bilingüismo. El volumen, publicado por Peter Lang en 2024, consta de siete capítulos, que abordan distintos aspectos relativos a los procesos cognitivos que subyacen al comportamiento lingüístico de hablantes que se comunican en dos o más lenguas. Este hecho, que en un primer momento se consideró un obstáculo para un buen desarrollo lingüístico (Goodenough, 1926), se valora en la actualidad como un interesante punto de partida para investigaciones empíricas, pues permite analizar asuntos tan variados como si hay dos sistemas lingüísticos independientes o interrelacionados (y, de ser esta última opción cierta, cómo y cuánto interactúan), qué estructuras neurofuncionales soportan todos estos procesos —y si difieren de las de un hablante monolingüe— o qué modelos de representación y acceso a la memoria a largo plazo pueden dar cuenta de los datos obtenidos en los trabajos experimentales. Estos interrogantes —y otros muchos— se plantean a lo largo del volumen.

Uno de los hilos conductores que articula todas las contribuciones es la importancia de considerar a cada hablante como único: en el caso que nos atañe, podemos decir que cada bilingüe lo es a su manera (esto es lo que se conoce como *la individualidad del bilingüe*; cf. Bruine, 2019): hay bilingües que lo son desde su infancia, porque cada progenitor les habló en una lengua distinta; hay otros que han aprendido una segunda lengua en un ambiente de inmersión, pero en edad más avanzada; e incluso hay hablantes cuya lengua materna puede perder dominancia

porque es la L2 la herramienta que utilizan de manera continua en su día a día (lo que encaja dentro de los procesos de *atrición*; cf. Dorado Escribano, 2020). Es, por tanto, esencial tener en cuenta las diferencias individuales a la hora de estudiar el sistema lingüístico. Pese a todas estas posibles variaciones, cualquier hablante bilingüe demuestra una gran capacidad en la gestión de dos sistemas lingüísticos en un único espacio representacional; es decir, en un único cerebro, como se refleja en el título del libro.

Dentro de esta capacidad de gestión de dos o más lenguas, se pueden encontrar ciertas regularidades, que son las que se exponen a lo largo de los distintos capítulos del monográfico. Y se hace apoyando los distintos planteamientos teóricos con trabajos experimentales, que permiten reforzarlos o rebatirlos. El empirismo es, por tanto, otro de los grandes protagonistas del volumen. En concreto, los capítulos se centran en el procesamiento del lenguaje y se enmarcan en una perspectiva psicolingüística. No obstante, tal y como comenta la propia editora en la introducción, el bilingüismo «no conforma un fenómeno homogéneo y monolítico» (p. 12), por lo que existen otros ámbitos de investigación que escapan al alcance del monográfico; por ejemplo, el bilingüismo estudiado desde una perspectiva sociolingüística —a la hora de atender fenómenos como la diglosia— o el bilingüismo analizado desde una óptica aplicada, como podría ser en el contexto educativo. No pretende esta observación consignar ningún demérito de la monografía reseñada, sino más bien recalcar una vez más la alta complejidad de este campo de estudio, aparentemente inabarcable.

Este viaje por la mente del bilingüe se inicia desde el nivel fonológico, con un capítulo en el que Miquel Llompart y Miquel Simonet tratan de responder a la pregunta de cómo procesa un bilingüe los sonidos de las lenguas que conocen. Para ello, se presentan distintos trabajos experimentales que han analizado cómo se procesan tanto las unidades segmentales y suprasegmentales como las propias palabras. Se hace mucho hincapié en la necesidad de estudiar el procesamiento fonológico en interacción con el léxico; este último es el protagonista del segundo capítulo, firmado por Montserrat Comesaña. Esta autora repasa los distintos modelos que han sido propuestos para dar cuenta de cómo los hablantes bilingües reconocen las unidades léxicas de una de sus lenguas (y de cómo inhiben las otras candidatas). El recorrido es exhaustivo y viene respaldado por ese ingrediente principal que dota de coherencia a todo el volumen: los datos experimentales, que, en este caso, apuntan a que ambas lenguas están siempre activadas y que el léxico parece procesarse de manera selectiva (es decir, atendiendo a variables puramente

lingüísticas, y no tanto contextuales). El siguiente nivel de procesamiento que se presenta es el de las oraciones. José Manuel Igoa inicia su capítulo poniendo de relieve que entender oraciones es una actividad compleja que cualquier hablante nativo hace sin esfuerzo. Esta aparente contradicción es todavía más sorprendente cuando añadimos el componente bilingüe. En este capítulo el autor dedica un apartado a analizar, de manera accesible, qué significa realmente comprender una oración, para después presentar los distintos modelos disponibles en relación con el procesamiento bilingüe.

Una vez que el lector se ha familiarizado con los distintos niveles de procesamiento, se pasa a comentar un asunto inseparable del bilingüismo: el cambio de código. Uno de los objetivos del capítulo firmado por Paola E. Dussias y Jessica Vélez-Avilés consiste en desmontar el mito de que el cambio de código es una muestra de pobreza lingüística (o de *alingüismo*); de hecho, argumentan que es una habilidad compleja que requiere de un alto nivel de competencia y en la que han de intervenir, por tanto, otras funciones ejecutivas que van más allá de lo lingüístico. Se admite generalmente que estas funciones ejecutivas constituyen uno de los factores que determinan las acusadas diferencias individuales. Por tal motivo son el objeto del siguiente capítulo, escrito por Alicia Luque e Irene Finestrat. Estas autoras presentan una taxonomía detallada de los factores aludidos, que son clasificados en internos y externos. Los internos, a su vez, se dividen entre los puramente lingüísticos —como la competencia lingüística en la L2— y los cognitivos, como la memoria procedimental. Los externos, por su parte, corresponden, por ejemplo, al contexto de adquisición o aprendizaje de las distintas lenguas implicadas.

El penúltimo capítulo del monográfico, firmado por Olga Ivanova, está dedicado a las «huellas cognitivas del bilingüismo» (p. 227), pues es innegable que nuestro cerebro está dotado para sostener el uso concurrente de más de una lengua. Además del recorrido por los distintos modelos dedicados a analizar la topografía cerebral y las conexiones neuronales, resultará especialmente interesante al lingüista la sección dedicada a la tipología de las lenguas, pues se plantea la pregunta de si hay diferencias cognitivas en función del tipo de lenguas que el bilingüe domine.

El volumen se cierra con un trabajo de Elena Garayzábal Heinze e Irene Hidalgo de la Guía, que estudia la relación entre bilingüismo y trastornos del lenguaje. En primer lugar, se centran en los trastornos que afectan a los adultos (como las afasias o la demencia), para pasar a revisar aquellos más frecuentes en la infancia (como el TDL o el TEA), analizando si hay diferencias entre un hablante monolingüe y uno bilingüe. Dedican también una sección a las lenguas de signos, que no han de ser olvidadas en el estudio lingüístico y cuyo análisis, además, supone poner

de manifiesto los principios de igualdad y dignidad de las lenguas, como afirma Moreno Cabrera (2016).

Espero que este breve recorrido haya convencido al lector a adentrarse en el volumen de Álvarez-García, que se puede considerar una referencia ineludible para cualquier persona interesada en los hablantes bilingües y su(s) sistema(s) lingüístico(s).

REFERENCIAS

- Bruine, A. de (2019). Not all bilinguals are the same: a call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences*, 9(3), 1-13.
- Dorado Escrivano, M. G. (2020). Atrición lingüística del español como primera lengua. *Lengua y Habla*, 24, 54-93.
- Goodenough, F. (1926). Racial difference in the intelligence of school children. *Journal of Experimental Psychology*, 9, 388-397.
- Moreno Cabrera, J. C. (2016). *La dignidad e igualdad de las lenguas. Crítica de la discriminación lingüística*. Madrid: Alianza Editorial.

Natalia LÓPEZ-CORTÉS
Universidad de Zaragoza

MOLINA MARTOS, I., HERNÁNDEZ, E., MARTÍN BUTRAGUEÑO, P., & MEDIETA, E. (eds.).
Caminos y palabras: Estudios de variación lingüística dedicados a Pilar García Mouton. Valencia: Tirant Humanidades, 2024, 994 pp.
https://doi.org/10.36151/TLB_9788411831598

Caminos y palabras: Estudios de variación lingüística dedicados a Pilar García Mouton es el título del volumen publicado en Tirant Lo Blanch en 2024 con el que más de cincuenta lingüistas de reconocido prestigio homenajearon a la académica con motivo de su jubilación. Este título alude a las palabras y a su vínculo con el mundo en el que se establece; caminos que se tornan posibilidades en la variación lingüística, que sirve como nexo temático. Así como los caminos recorridos haciendo encuestas dialectales. Molina Martos, Hernández, Martín Butragueño y Mendieta, los cuatro editores, así lo escriben en su introducción: «Entendemos, eso sí, que lo que vincula mayormente a las investigaciones de variación es la relación entre lenguaje (las palabras) y los hechos históricos, sociales y culturales (las cosas)».

Si supone todo un reto articular una obra coral tan extensa sin reducirla a una mera acumulación de artículos, en este caso, ese propósito se ha cumplido.

La contribución de Pilar García Mouton a la filología moderna está marcada por su curiosidad y su respeto por la diversidad, vista como fuente de riqueza, tal y como queda de manifiesto en el apartado del libro dedicado a su semblanza. En dicha sección, Isabel Molina hace un repaso de la trayectoria académica de García Mouton desde que comenzó a estudiar Filología Románica. Además de por su excelente labor en el ámbito de la geografía lingüística, en el que es uno de los exponentes hispanos más importantes, la homenajeada ha destacado en el estudio del habla de las mujeres, al que se ha aproximado en publicaciones como *¿Cómo hablan las mujeres?* (Arco Libros, 1999) o *Así hablan las mujeres. Curiosidades y tópicos del uso femenino del lenguaje* (La Esfera de los Libros, 2003); en la edición y publicación de textos académicos, tanto en la vertiente editorial con Gredos, como en la prestigiosa *Revista de Filología Española*; en la divulgación en medios de comunicación con programas como *Palabras Moribundas* y en la gestión académica, pues cabe recordar que ha dirigido el Centro de Humanidades (2003-2005), el Instituto de la Lengua Española (2002-2008) y el Instituto de Lengua, Literatura y Antropología. Ha sido vicerrectora de Enseñanzas del Español y otras lenguas en la Universidad Internacional Menéndez Pelayo (2019-2021) y coordinadora del Área de Filología y Filosofía de la Agencia Nacional de Evaluación y Prospectiva (2005-2008). García Mouton es una autoridad que ha configurado su praxis investigadora con sensibilidad ante los márgenes lingüísticos y sociales como espacios legítimos desde los que ampliar las fronteras del conocimiento.

Este volumen, que supera las novecientas páginas, gira en torno a un eje común: la variación lingüística, abordada desde múltiples perspectivas y con una mirada que trasciende el ámbito del español. En sus páginas se refleja tanto la calidad científica como el compromiso humano de sus autoras y autores, cuyas contribuciones se extienden también a otras lenguas como el italiano, el gallego, el portugués, el rumano, el catalán y el valenciano o el vasco. Conviven aquí distintas generaciones de especialistas que, desde enfoques diversos, enriquecen el estudio de la variación lingüística. La obra se estructura en varios bloques: una introducción a cargo de los editores, donde se presenta y justifica el hilo conductor temático y la organización interna del volumen; una semblanza biográfica de Pilar García Mouton, firmada por Molina Martos, que recorre con detalle su trayectoria profesional y vital; y cinco secciones temáticas que recogen cuarenta y seis aportaciones donde se estudia la variación lingüística desde diferentes enfoques.

Las dos primeras secciones ponen el léxico en el centro. En la primera de ellas, (1) «El léxico en su geografía», se agrupan diez trabajos que ponen el foco en unidades léxicas estudiadas desde su dimensión espacial, además de otro conjunto de aportaciones que examinan situaciones de lenguas en contacto también desde su dimensión diatópica. En el apartado (2), «El léxico: a vueltas con su propia historia», los autores abordan la lingüística histórica y el cambio lingüístico a partir de los procesos neológicos, ya sea mediante la adopción o la creación de palabras, entendidos como mecanismos habituales tanto de formación léxica como de desarrollo de significados concretos, ya mediante su proyección en los diccionarios. Tras estos diferentes enfoques subyace una misma idea formulada por los editores del volumen en la introducción: «en el terreno de la cultura material, natural o intelectual, la historia de las palabras normalmente coincide con la historia de las cosas». En esta sección, los trabajos se sitúan en la tradición de los diccionarios filológicos y en la romanística.

El tercer apartado del volumen, titulado (3) «Fuerzas analógicas: el papel de la morfología y la sintaxis», reúne nueve contribuciones centradas en un amplio abanico de fenómenos gramaticales, vertebrados en torno a varios ejes comunes que les confieren coherencia interna. Uno de los principales puntos de conexión de estos trabajos es el uso compartido de datos procedentes de distintas fuentes de cartografía lingüística, lo que permite abordar de manera empírica diversos aspectos de la variación morfosintáctica. Este bloque parte del reconocimiento de que los procesos analógicos —entendidos en un sentido amplio— operan como un motor de cambio gramatical, especialmente en su papel de mecanismo de regularización y emparejamiento entre formas disímiles, lo que da lugar a la formación de paradigmas. Estas fuerzas analógicas, que se intuyen desde la tradición gramatical más temprana, comienzan a sistematizarse en el siglo XIX y experimentan un notable desarrollo teórico a partir del último cuarto del siglo XX, cuando se amplía su estudio no solo desde una perspectiva formal o semántica, sino también cognitiva. Asimismo, se pone de relieve que la evolución del análisis de la variación morfosintáctica ha requerido un proceso más lento y complejo que el de la variación fónica, en tanto que muchas de las alternancias gramaticales implican matices semánticos y pragmáticos entre formas alternativas.

En la sección (4) del libro, titulada «El impulso de la regularidad: la dimensión fónica», se reúnen trabajos dedicados al análisis lingüístico del componente sonoro y oral de las lenguas. Los trabajos parten en muchos casos de materiales extraídos de atlas lingüísticos y de muestreos en comunidades de habla, recursos

que tradicionalmente han resultado especialmente productivos para el estudio de la variación fónica segmental en dialectología y sociolingüística. A ello se suma, de manera más reciente, el interés por el análisis de fenómenos prosódicos desde la perspectiva de la variación, lo que aporta una dimensión adicional al enfoque sonoro.

La última sección del volumen, titulada (5) «Variación, filología y sociedad», agrupa un conjunto de contribuciones que, aunque heterogéneas en sus aproximaciones y objetos de estudio, comparten una voluntad común de explorar la variación lingüística desde perspectivas complementarias. Algunas de ellas se enmarcan dentro del ámbito de la filología tradicional, mientras que otras adoptan enfoques teóricos más actuales o trabajan con corpus distintos a los empleados en los apartados anteriores. Pese a esta diversidad metodológica y temática, los trabajos presentan una coherencia interna que refuerza la estructura del volumen en su conjunto y permite cerrar el recorrido con una visión amplia y enriquecida sobre los múltiples modos de abordar el cambio y la variación en las lenguas. Esta obra se configura así como un punto de referencia para futuras investigaciones en el campo de la variación lingüística, ofreciendo un compendio representativo de las principales líneas de trabajo que se desarrollan en la actualidad.

Esta es una obra dotada con una gran carga emotiva, que va de lo académico a lo humano, y se convierte en un conjunto de cartas de agradecimiento en el que los autores aprovechan para remarcar la influencia que ha tenido García Mouton en su propio camino. La pluralidad de enfoques que reúne *Caminos y palabras* es, sin duda, uno de los mayores aciertos del volumen. Más de cincuenta autores, de diferentes instituciones y trayectorias, aportan su mirada para trazar un mapa diverso. Además, el volumen representa a la homenajeada. Cada contribución evoca o dialoga con sus líneas de investigación. Queda reflejado así el vínculo personal, la admiración profesional y también el rigor académico.

Caminos y palabras celebra a Pilar García Mouton y nos recuerda que la lingüística es un campo en movimiento, en diálogo constante con la sociedad. Lo mejor de las palabras no es lo que fijan, sino lo que abren: caminos y posibilidades.

Gema CUESTA
Universidad de Alcalá

Edita
sel

<http://revista.sel.edu.es>

