

SALVADOR PONS BORDERÍA, ADRIÀ PARDO LLIBRER
Y AMPARO ALEMANY MARTÍNEZ

LA MARCACIÓN DISCURSIVA EN ESPAÑOL:
DESCRIPCIÓN Y ANÁLISIS ESTADÍSTICO
DESDE EL *DPDE*

 EDITORIAL
UNIVERSIDAD DE SEVILLA

Sevilla 2023

Colección Lingüística
Núm.: 85

COMITÉ EDITORIAL:

Araceli López Serena
(Directora de la Editorial Universidad de Sevilla)
Elena Leal Abad
(Subdirectora)

Concepción Barrero Rodríguez
Rafael Fernández Chacón
María Gracia García Martín
María del Pópulo Pablo-Romero Gil-Delgado
Manuel Padilla Cruz
Marta Palenque
María Eugenia Petit-Breuilh Sepúlveda
Marina Ramos Serrano
José-Leonardo Ruiz Sánchez
Antonio Tejedor Cabrera

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética o cualquier almacenamiento de información y sistema de recuperación, sin permiso escrito de la Editorial Universidad de Sevilla.

Este trabajo ha sido posible gracias al proyecto CIPROM/2021/038 «Hacia la caracterización diacrónica del siglo XX» (DIA20), de la Generalitat Valenciana, y al proyecto de I+D+i PID2021-125222NB-I00 «Aportaciones para una caracterización diacrónica del siglo XX», financiado por MCIN/ AEI /10.13039/501100011033/ y por FEDER 'Una manera de hacer Europa'.



© Editorial Universidad de Sevilla 2023
Porvenir, 27 - 41013 Sevilla
Tlfs.: 954 487 447; 954 487 451; Fax: 954 487 443
Correo electrónico: info-eus@us.es
Web: <https://editorial.us.es>

© Salvador Pons Bordería, Adrià Pardo Llibrer y Amparo Alemany Martínez 2023

Impreso en papel ecológico
Impreso en España-Printed in Spain

ISBN: 978-84-472-2512-5
Depósito Legal: SE 2372-2023

Diseño de cubierta: Notanumber
Maquetación y realización de cubierta: Dosgraphic s.l. (dosgraphic@dosgraphic.es)
Impresión: Masquelibros

Índice

CAPÍTULO 1.	
INTRODUCCIÓN.....	11
1.1. Presentación y objetivos.....	11
1.2. El <i>DPDE</i> como base de datos para el estudio de la marcación discursiva	13
1.2.1. Un corpus heterogéneo: tipo, muestra, variación diatópica y diafásica.....	13
1.2.2. Homogeneización de las categorías y homogeneización de las funciones	16
1.2.2.1. De las partículas a los marcadores: un camino de ida y vuelta.....	16
1.2.2.2. Delimitación y homogeneización de funciones pragmáticas	18
1.3. El modelo Val.Es.Co. para la segmentación en unidades discursivas	22
1.3.1. Un modelo de unidades jerárquico.....	23
1.3.2. El modelo en el discurso escrito	28
1.3.3. Un sistema de posiciones	31
1.4. Para un análisis estadístico de la marcación discursiva	33
1.4.1. Segmentación y obtención de datos.....	34
1.4.2. Hacia la cuantificación de las variables.....	36
CAPÍTULO 2.	
ESTADÍSTICA DESCRIPTIVA, PRUEBAS CHI-CUADRADO Y DISTRIBUCIONES BINOMIALES.....	41
2.1. Combinaciones de la variable <i>registro</i>	42
2.1.1. Registro por unidad	42
2.1.2. Registro por función general.....	46
2.1.3. Registro por unidad por función general.....	48
2.1.4. Registro por unidad por posición	49
2.1.5. Registro por posición.....	51

2.2. Combinaciones de la variable <i>función general</i>	53
2.2.1. Función general por posición.....	54
2.2.1.1. Función general por posición (registro formal)	55
2.2.1.2. Función general por posición (registro informal).....	56
2.2.2. Función general por unidad	58
2.2.2.1. Función general por unidad (registro formal)	60
2.2.2.2. Función general por unidad (registro informal).....	61
2.2.3. Función general por unidad-posición.....	63
2.3. Combinaciones de la variable <i>función específica</i>	68
2.3.1. Función general por función específica	69
2.3.2. Función específica por registro.....	70
2.3.3. Función específica por función general por registro.....	71
2.3.4. Función específica por posición	72
2.3.5. Función específica por unidad.....	77
2.3.6. Función específica por unidad por posición.....	79
2.4. Combinaciones de la variable <i>marcador</i>	80
2.4.1. Marcadores por unidad y registro (formal e informal, respectivamente).....	80
2.4.2. Marcadores por unidad y posición.....	89
2.4.3. Marcadores por registro	98
2.4.4. Marcadores por función general por función específica..	99
2.5. Conclusiones	101
CAPÍTULO 3.	
LOS MARCADORES DEL DISCURSO DESDE UNA PERSPECTIVA MULTIVARIANTE	103
3.1. Analizando las funciones discursivas	104
3.1.1. Introducción	104
3.1.2. Análisis factorial: pasos y datos.....	107
3.1.3. Análisis factorial: datos e interpretación	109
3.1.4. Conclusiones	119
3.2. Análisis de tipo clúster	119
3.2.1. Clúster aglomerativo jerárquico (CAJ)	121
3.2.2. Clúster de K-medias	130
3.2.3. Afinando el análisis: valores extremos y partículas no marcadores.....	141

CAPÍTULO 4.	
CONCLUSIONES	149
4.1. Cuestiones metodológicas.....	149
4.1.1. Valor del material lexicográfico en el estudio de los marcadores discursivos	149
4.1.2. Capacidad explicativa del modelo Val.Es.Co.....	150
4.1.3. Aplicación de la estadística a la investigación pragmática	153
4.2. Cuestiones teóricas.....	155
4.2.1. Aspectos sistemáticos en la descripción de la marcación discursiva en español	155
4.2.2. Investigación futura	162

APÉNDICES

APÉNDICE 1. LISTADO DE ABREVIATURAS	167
APÉNDICE 2. GLOSARIO DE TÉRMINOS ESTADÍSTICOS.....	169
APÉNDICE 3. GLOSARIO DE TÉRMINOS LINGÜÍSTICOS.....	175
APÉNDICE 4. SIGNOS Y CONVENCIONES DE TRANSCRIPCIÓN DEL GRUPO VAL.ES.CO.....	183
APÉNDICE 5. NOTACIÓN PARA LA SEGMENTACIÓN EN UNIDADES DISCURSIVAS	185
BIBLIOGRAFÍA.....	187

CAPÍTULO 1. INTRODUCCIÓN

1.1. Presentación y objetivos

El presente trabajo es un estudio global y cuantitativo de la marcación discursiva en español, realizado a partir del comportamiento de la clase funcional de palabras que la realiza: los marcadores del discurso.

En la bibliografía del español, son numerosos los estudios cualitativos sobre un marcador (de los que dan cuenta trabajos conjuntos como Loureda y Acín-Villa 2010 o Borreguero Zuloaga y Pons Rodríguez en prensa), así como los estudios cualitativos sobre grupos de marcadores (Fuentes Rodríguez 1987, Mederos Martín 1988, Portolés 1998, Martín Zorraquino y Portolés 1999, por citar solo algunas referencias clásicas). También existen compendios sobre experimentos diseñados para medir las estrategias de procesamiento de los marcadores (Loureda *et al.* 2022), pero –salvo error u omisión– no existe un estudio conjunto de corte cuantitativo sobre el funcionamiento de esta clase de palabras, en el que se respondan las preguntas relativas a sus funciones más frecuentes, a sus posiciones preferidas, a los parámetros más relevantes para su análisis o a la relación existente entre los diferentes planos del discurso, tanto oral como escrito. Este trabajo avanza una respuesta a estas y otras cuestiones. En un estudio de este tipo, tan importante es el *cuánto* como el *cómo*, es decir, cuál ha sido el proceso de selección y ordenación de los datos analizados, ya que dicho proceso determina en buena medida los resultados; en este primer capítulo, se explica cómo se han abordado tales cuestiones.

Desde el punto de vista teórico, el presente estudio parte del marco metodológico propuesto en el modelo de unidades del grupo Val.Es.Co. (Briz y Grupo Val.Es.Co. 2003, Grupo Val.Es.Co. 2014, Pons Bordería 2022). En dicho modelo, se considera que la función de un marcador viene predeterminada por la interrelación entre la unidad discursiva en la que aparece –un segmento delimitado del discurso hablado– y la posición que ocupa con respecto a dicha unidad (*vid.* § 1.3 para una explicación detallada). La ventaja de este modelo es que ofrece una serie de rasgos que se pueden aplicar homogéneamente a todos los marcadores estudiados: unidad, posición, registro y función.

Desde el punto de vista operativo, los datos del análisis se han extraído del *Diccionario de partículas discursivas del español* (Briz, Pons y Portolés 2008; en adelante, *DPDE*), en el que se describe el comportamiento pragmático de más de un centenar de partículas discursivas del español. La figura 1 reproduce una de estas entradas:

The screenshot shows the entry for 'en cambio' in the 'Diccionario de partículas discursivas del español'. The page is titled 'en cambio' and is by José Portolés. It includes a search bar with 'en cambio' entered, a navigation menu with 'INTRODUCCIÓN', 'ENTRADAS', and 'REDACTORES', and a search icon. The main content area is divided into sections: 'DEFINICIÓN', 'MAS EJEMPLOS', and a table of 'PUNTO DE COMPARACIÓN' and 'CONTRASTE'. The 'DEFINICIÓN' section states: 'Indica que el miembro del discurso en el que aparece contrasta con un miembro discursivo anterior:'. It provides an example: 'A Prullás aquel tabaco tan áspero, en ayunas, le producía náuseas. Don Lorenzo Verdugones, en cambio, se mostraba la mar de satisfecho.' and cites 'E. Mendoza, *Una comedia ligera*, Barcelona, Seix Barral, 1996, 307'. The 'MAS EJEMPLOS' section includes '1. oral' with the example: '¿Cómo recuerda aquella época? Porque nos la pintaron muy bien, en cambio parece que habla componentes angustiosos, ¿no?' and cites 'Oral, España, CREA, 1991'. The 'CONTRASTE' table shows: 'A Prullás' vs 'aquel tabaco le producía náuseas' and 'Don Lorenzo Verdugones' vs 'en cambio, se mostraba la mar de satisfecho'. The bottom of the page has the text 'Diccionario de Partículas Discursivas del Español'.

Figura 1. Ejemplo de entrada del *DPDE* (<http://www.dpde.es/#/entry/encambio>)

Cada entrada caracteriza discursivamente una partícula; en su mayoría, marcadores, pero también conectores y otros operadores (*vid.* § 1.2.2). Con un metalenguaje accesible para un usuario no especialista, el *DPDE* es una obra lexicográfica que ha contribuido a la descripción sistemática de los marcadores del discurso en español, gracias al esfuerzo conjunto de la cuarentena de autores que han colaborado en la realización del diccionario.

Tomado en su conjunto, el *DPDE* puede considerarse como una gran base de datos descriptiva susceptible de analizarse de modo cuantitativo; sin embargo, dado que refleja los diferentes enfoques de cada uno de los investigadores que han contribuido a ella, es necesario, para una descripción unitaria, homogeneizar previamente los datos del *DPDE*. Tras este trabajo previo, los datos pueden cuantificarse, estudiarse y compararse.

A continuación, se explican los pasos seguidos para unificar, seleccionar y analizar los datos. En primer lugar (§ 1.2), se trata el problema de la heterogeneidad del *DPDE* como base de datos; esto es, la variedad de la muestra en la que se basa cada descripción (§ 1.2.1); para poder analizarla de manera unitaria (§ 1.2.2), es imprescindible tener en cuenta las diferencias categoriales (en nuestro caso, marcadores

discursivos como objeto de estudio frente a otras *partículas*; § 1.2.2.1), así como las diferencias funcionales (el hecho de que una misma función se describa como *acuerdo*, *aceptación* o *validación* en diferentes entradas; § 1.2.2.2). Tras esta doble delimitación, se presenta el modelo Val.Es.Co. para la segmentación discursiva (§ 1.3), que propone un análisis sistemático de los marcadores en términos de unidades –tanto en el dominio oral (§ 1.3.1) como en el escrito (§ 1.3.2)– y posiciones (§ 1.3.3). Finalmente (§ 1.4), se expone el proceso seguido en la sistematización de las entradas (§ 1.4.1) y se introducen las cuestiones previas a su análisis cuantitativo (§ 1.4.2).

1.2. El DPDE como base de datos para el estudio de la marcación discursiva

Bajo el término genérico de *partícula*, el DPDE recoge una amplia variedad de formas. La descripción varía de una entrada a otra: no es posible describir del mismo modo elementos propios de la conversación coloquial (como *oye* o *¿eh?*) que elementos que introducen escalas (*hasta*, *incluso*) o dependen de un ámbito negativo previo (*menos aún*). Aunque todas las entradas del DPDE siguen las mismas instrucciones de macroestructura y microestructura, es necesario solventar ciertos problemas propios de la información lexicográfica para extraer de esta un conjunto homogéneo de datos. Estos problemas son los siguientes: por una parte (§ 1.2.1), la relación entre tipos y muestras y la consiguiente desproporción entre registros y dialectos y, por otra parte (§ 1.2.2), la definición de la categoría *marcador discursiva* (§ 1.2.2.1), así como de las funciones pragmáticas que esta categoría puede desempeñar (§ 1.2.2.2).

1.2.1. Un corpus heterogéneo: tipo, muestra, variación diatópica y diafásica

Cada una de las entradas del DPDE presenta una serie de ejemplos que ilustran los distintos usos de un marcador. Se trata de casos tomados de diferentes fuentes y utilizados para la ejemplificación lexicográfica. Esta selección de ejemplos es enteramente cualitativa: depende en gran medida de los conocimientos del redactor y no se ciñe necesariamente a criterios de representatividad. Sin embargo, la macroestructura del diccionario permite entender los ejemplos de cada entrada como usos prototípicos del marcador en cuestión. Por ejemplo, en el caso de *pues*:

pues ¹		MARGARITA PORROCHE BALLESTEROS
OTROS USOS	<p>Aparece en intervenciones reactivas o réplicas que no son contestaciones a preguntas:</p> <p>A: pero yo lo hacía sin- sin conocimiento§ B: § de causa/ y de efecto (RISAS). A: pues tú lo mismo/ así que no hables</p> <p>A. Briz y Grupo Val.Es.Co. <i>Corpus de conversaciones coloquiales</i>, Madrid, Arco Libros, 2002, 65, 1. 609-612</p> <p>En estos casos, además de resaltar una información nueva, indica una relación entre los enunciados que une basada en una cierta oposición marcada por un cambio en la línea de argumentación (por ejemplo, el hablante muestra su desacuerdo ante las palabras de su interlocutor o una crítica) o un contraste entre modalidades (este tipo de pues aparece frecuentemente introduciendo enunciados exhortativos, exclamativos, interrogativos, etc.). Obsérvese que, ante el enunciado <i>La última película de Almodóvar es estupenda, puede decirse ¡Pues a mí no me gusta! / Pues vete a verla solo / Pues la ponen en El Dorado, pero no son esperables respuestas del tipo ¡Pues a mí me gusta! o ¡Pues a mí Almodóvar me parece un director magnífico!</i></p> <p>Pues aparece también como primer elemento de un acto de comunicación:</p> <p>Pues venía a pedirle un favor.</p> <p>Sirve para enlazar circunstancias implícitas en el acto comunicativo con el miembro del discurso que sigue a pues. Circunstancias como «usted se preguntará para qué he venido», pues (información nueva) venía a pedirle un favor.</p>	

Figura 2. Funciones del marcador en el campo *OTROS USOS* de la entrada para *pues*

pues ¹		MARGARITA PORROCHE BALLESTEROS
SINTAXIS	<p>Pues puede ir seguido de oraciones y de todos los tipos de sintagmas no oracionales. El único requisito es que el elemento que introduzca transmita información nueva. En el ejemplo que sigue introduce un sintagma preposicional:</p> <p>Yo pienso que, vaya, hay claras evidencias de que no es lo mismo un drogadicto en una clase media, en una clase alta, que un drogadicto pues en el Pozo del Tío Raimundo.</p> <p>Oral, España, CREA, 1983</p> <p>Puede también encabezar la proposición principal —cuando esta es una información nueva— de una oración compuesta:</p> <p>C: cuando tenga que ir ahí al lao del tren/ pues lo llevaré yo</p> <p>A. Briz y Grupo Val.Es.Co. <i>Corpus de conversaciones coloquiales</i>, Madrid, Arco Libros, 2002, 256, 1.504</p> <p>O introducir la oración subordinada o una oración coordinada e, incluso, una oración de relativo. Pues precede a cualquier elemento que pueda expresar una información nueva:</p> <p>Es bonito, pues porque coges un mueble que está muy viejo, todo pintao o hecho polvo y entonces, pues lo limpias bien limpio y lo dejas en condiciones.</p> <p>Corpus de lengua hablada de Zaragoza, 1988</p> <p>Por otra parte, pues puede ir precedido y seguido de elementos incidentales que están fuera de la estructura informativa del enunciado.</p>	

Figura 3. Funcionamiento del marcador en el campo *SINTAXIS* de la entrada para *pues*

Los campos *DEFINICIÓN* y *OTROS USOS* (figura 2) recogen las principales funciones de *pues*, mientras que los campos *POSICIÓN* y *SINTAXIS* (figura 3) añaden información sobre el lugar que ocupa el marcador en el enunciado, que puede ser el mismo o cambiar de una función a otra. Desde el punto de vista cuantitativo, los ejemplos de cada entrada del *DPDE* deben verse como *tipos* que reproducen los posibles usos de un marcador; son explicativos cualitativamente, pero no pretenden constituirse en la muestra representativa de una población (puesto que se persigue la observación sistemática de las funciones de un marcador, no su mayor o menor presencia el corpus).

El registro es otro de los factores que hacen del *DPDE* un corpus heterogéneo. El *DPDE* organiza el campo *REGISTRO* bajo las nociones *formal* e *informal*:

pues ¹		MARGARITA PORROCHE BALLESTEROS
REGISTRO	Está marcado por su uso frecuente en el registro coloquial y en el discurso oral. Cuando aparece en el escrito, supone una reproducción o una imitación del oral.	
Diccionario de Partículas Discursivas del Español		

Figura 4. Campo *REGISTRO* en relación con las etiquetas de *formal/escrito* e *informal/coloquial*

Formal e *informal* no se distribuyen de manera equilibrada en el *DPDE*. Hay que tener en cuenta que la mayoría de los marcadores son propios de la escritura y, por tanto, del registro formal. Frente a este grupo mayoritario, una pequeña parte de los marcadores en el *DPDE* es exclusiva del registro informal conversacional. En este grupo, los ejemplos de cada entrada están tomados de fuentes orales (conversaciones coloquiales) o bien de textos escritos que reproducen la oralidad (como el teatro o las novelas). El tema de la mimesis de la oralidad se ha trabajado con profusión (López Serena 2007, 2009, Grupo Val.Es.Co. 2022), al igual que se han estudiado géneros orales que reproducen las condiciones de lo formal, como entrevistas o conferencias (Briz 2010). Más que de lo oral y lo escrito, que no son sino formas de realización, se trata aquí de la oralidad y la escrituralidad (Kabattek y Jacob 2001). Para los objetivos del presente trabajo, que no es sino el análisis conjunto de la marcación discursiva en español, es necesario tener en cuenta el diferente peso de estos dos grupos en el análisis y separar la muestra por registros cuando sea conveniente.

En cuanto a la variación diatópica, aunque la descripción tome como base el español peninsular, se han incluido también las entradas con marcadores propios del español de América (figura 5):

ándale ¹		MARIA EUGENIA FLORES TREVIÑO
DEFINICIÓN	DISTRIBUCIÓN GEOGRÁFICA	
	México	
	<p>Ándale (ánde) Apela con fuerza al interlocutor para que lleve a cabo una acción, ya sea a modo de orden o de invitación:</p> <p>Pero sí, ándale, hija, tráigame leña. Allí tráigame agua. Ándele. Lave el traste, ándele. Quiebra el nixtamal, ándele.</p>	

Figura 5. Etiqueta *Distribución geográfica* en los marcadores del español americano

Si bien solo se indica la distribución geográfica en los casos del español americano, los marcadores cuya descripción es válida para ambos lados del Atlántico

constituyen el grupo mayoritario¹. En este sentido, el estado actual de las investigaciones impide por ahora el estudio cuantitativo de la variación diatópica de la marcación discursiva, pero no una primera descripción general del fenómeno para el diasistema del español. Investigaciones posteriores complementarán los datos este primer estudio.

1.2.2. Homogeneización de las categorías y homogeneización de las funciones

La cuantificación unitaria de los datos del *DPDE* presenta dos problemas adicionales: ni todas las formas recogidas en el *DPDE* son marcadores, aunque puedan compartirse ciertas intuiciones sobre su funcionamiento, ni la descripción de las funciones que desempeña cada partícula se ajusta a un número finito de etiquetas, pues se mantiene la explicación de cada autor en sus propios términos. Para solucionar estos dos problemas, se hace necesaria una labor previa de discriminación y de estandarización. En primer lugar (§ 1.2.2.1), se parte de una definición operativa de *marcador discursivo* que permita discriminar estas formas de otras igualmente recogidas bajo la noción –más amplia– de *partícula discursiva*. En segundo lugar (§ 1.2.2.2), las funciones mediante las que se describe cada marcador no están homogeneizadas en el *DPDE*, por lo que es necesario, para su estudio cuantitativo, transformar dicha heterogeneidad en una nómina cerrada de funciones pragmáticas (§ 1.2.2.2).

1.2.2.1. De las partículas a los marcadores: un camino de ida y vuelta

El *DPDE* recoge diversas formas con una característica común: todas ellas conllevan algún tipo de repercusión discursiva en el enunciado en el que aparecen. Ahora bien, el hiperónimo *partícula discursiva* engloba un conjunto particularmente heterogéneo de categorías. Como se indica en la misma introducción del diccionario:

Constituye un logro importante haber unificado la terminología en torno a estas marcas del discurso, dada su heterogeneidad, y haber minimizado las diferencias conceptuales y teóricas entre los investigadores en lo que se entiende por «marcador», «conector», «partícula», «enlace textual» o «locución» (Briz, Pons y Portolés 2008: § 1.1).

1. Por ejemplo, al igual que la entrada de *al contrario* (redactada por J. Portolés, de la Universidad Complutense), la entrada de *en última instancia* (redactada por P. C. Hernández y L. Miñones, de la Universidad de Buenos Aires) no indica distribución geográfica, no porque sean marcadores propios del español de España, sino porque su descripción se aplica a todo el dominio hispánico (en línea con la noción estructuralista de diasistema; cf. Coseriu 1978, 1981).