

**EL CORPUS ORAL Y SONORO DEL ESPAÑOL RURAL (COSER) Y SU
CONTRIBUCIÓN AL ESTUDIO DE LA VARIACIÓN GRAMATICAL
DEL ESPAÑOL***

*Audible Corpus of Spoken Rural Spanish (COSER) and its contribution to the
study of grammatical variation in Spanish*

INÉS FERNÁNDEZ-ORDÓÑEZ
Universidad Autónoma de Madrid

ENRIQUE PATO
Université de Montréal

Resumen

El objetivo principal de este trabajo es dar a conocer la contribución pionera del COSER al estudio de la variación gramatical del español europeo y actualizar los informes previos. Su estructura es la siguiente: tras una breve introducción, se lleva a cabo una descripción general del corpus, así como de la metodología empleada en su elaboración. Después, se da a conocer el estado del COSER en línea (enero de 2019). Posteriormente, en el apartado principal del estudio, se resumen y sintetizan las aportaciones concretas que se han realizado hasta la fecha sobre numerosos aspectos de variación gramatical del español gracias a los datos del COSER (morfología verbal, concordancia de caso en los clíticos de objeto, concordancias de género y número, construcciones reflexivas, medias e impersonales con *se*, selección del modo, expresión del complementante, entre otros).

Palabras clave: español europeo, dialectología, variación gramatical, corpus lingüísticos, COSER

Abstract

The main objective of this work is to show COSER's pioneering contribution to the study of grammatical variation in European Spanish and to update previous reports. Its structure is as follows: after a brief introduction, a general description of this corpus is carried out, as well as the

* Este trabajo se enmarca dentro del proyecto "Cambios en el habla rural: del siglo XX (Atlas Lingüístico de la Península Ibérica, ALPI) al siglo XXI (Corpus Oral y Sonoro del Español Rural, COSER)", financiado por el Ministerio de Ciencia, Innovación y Universidades (PGC2018-095077-B-C2). El *Principio del deber contraído* (Labov) y el *Principio de gratitud lingüística* (Wolfram) establecen que todo investigador que disponga de datos lingüísticos de alguna comunidad de habla está en la obligación de ponerlos en conocimiento de dicha comunidad, y debe además buscar los modos prácticos de devolver el favor lingüístico proporcionado. Este trabajo es uno de esos modos para agradecer a todos los informantes su participación en las encuestas del COSER.

methodology used in its elaboration. Then, the status of COSER online is released so far (January 2019). Subsequently, in the main section of the paper, we summarize and synthesize all the contributions that have been made on numerous issues of grammatical variation in Spanish thanks to COSER data (verb morphology, case selection in object clitics, gender and number agreement, *se* constructions -reflexive, middle, and impersonal- mood selection, and complementizer expression, among others).

Keywords: European Spanish, dialectology, grammatical variation, linguistic corpora, COSER

1. INTRODUCCIÓN

Como es bien sabido, el estudio de la variación dialectal del español hablado en España (tanto en la península como en las islas) se ha basado principalmente en los datos de los diversos atlas lingüísticos regionales publicados y en algunas monografías de distinto valor e interés hasta hace relativamente poco tiempo (*cf.* Fernández-Ordóñez 2007c y ss.). En ambos tipos de trabajos la variación fonética y la variación léxica copaban casi toda la atención. Al hecho de que las monografías dialectales apenas estudiaban la variación gramatical debe sumarse el problema de que se habían centrado sobre todo en los dialectos llamados “históricos” (navarro, aragonés y asturleonés), sin atender en la misma medida al castellano rural. Esta situación deficitaria empezó a cambiar radicalmente con la creación –hace ya treinta años– del *Corpus Oral y Sonoro del Español Rural* (COSER), corpus de grabaciones del habla rural que empezó a compilarse en el año 1990 para complementar, precisamente, las fuentes tradicionales de información y los datos dialectales recopilados mediante cuestionarios no grabados, en especial, en lo relativo a la variación gramatical. Desde entonces, el COSER ha ido creciendo cada año, con sucesivas campañas de encuesta en la península, en Canarias y en Baleares, tal y como veremos más adelante y se puede comprobar en su página web (corpusrural.es).

Por otro lado, las fuentes escritas objeto de estudio en los trabajos sincrónicos solían reflejar solo la lengua estándar castellana, o el estándar regional de cada provincia, en el mejor de los casos. Por ese motivo, había toda una serie de fenómenos de variación gramatical que no se habían documentado en los corpus de referencia anteriores; y en el caso de que se encontraran, podía suceder que su localización geográfica fuera insegura, o que no se documentaran todos los contextos pertinentes para su completa descripción y análisis. No menos importante es el hecho de que la lengua oral –la de todos los sociolectos– tiene todavía poca presencia en las fuentes habituales (CREA, CORPES XXI) que los investigadores emplean en sus trabajos.

Las primeras campañas de encuesta realizadas en el marco del COSER siguieron de cerca los modelos de trabajo de Diego Catalán para el estudio de los romances tradicionales (cuyo fruto sería el *Archivo Sonoro del Romancero ‘Débora Catalán’*) y de Flora Klein-Andreu para el estudio de los clíticos *le/s*, *la/s*, *lo/s* en Castilla (la Vieja). Las grabaciones semidirigidas del COSER pronto mostraron que el trabajo de campo ofrece numerosísimas ventajas para la descripción y la explicación de los fenómenos gramaticales, al identificar por primera vez el trazado de la isoglosa de un rasgo determinado o al modificar y precisar una distribución geográfica ya conocida previamente. También sirven para desarrollar y evaluar nuevas hipótesis y explicaciones sobre aspectos de la variación gramatical, pues la cantidad y la calidad de los datos recopilados es mucho mayor de lo que se disponía previamente y, por tanto, su descripción lingüística más precisa. Las encuestas incluso proporcionan información gramatical desconocida, o muy poco estudiada hasta la fecha, por lo que se han podido

descubrir y describir nuevos fenómenos de variación morfosintáctica. De este modo, se han podido entender mejor numerosas estructuras dialectales del español europeo en sus distintas modalidades, septentrional, meridional, occidental y oriental, así como del español en contacto con otras lenguas peninsulares (§ 4).

Por otro lado, y como también tendremos ocasión de mostrar, los datos del COSER permiten comparar mejor los fenómenos registrados en el habla rural con aquellos propios del habla urbana, recopilados en los varios proyectos efectuados sobre la norma culta de las diversas ciudades españolas e hispanoamericanas, como el estudio sociolingüístico del español de España y de América (PRESEEA), o sobre el habla coloquial de algunas ciudades españolas (*Corpus Val.Es.Co* en Valencia, *Vernáculo Urbano Malagueño* en Málaga, *Corpus Oral de Conversa Col·loquial* en Barcelona, entre otros muchos). En efecto, la uniformidad en la metodología empleada en el COSER lo hace útil para medir tanto la distancia lingüística que separa diversas áreas (distancia geográfica) como la distancia lingüística que separa el grupo social rural de mayor edad de otros como el de los hablantes de mayor nivel sociocultural o el de los hablantes más jóvenes (distancia sociológica)¹.

Por último, los datos del COSER también permiten valorar en su justa medida la evolución diacrónica de ciertos fenómenos gramaticales al ser comparables con otras fuentes previas de la dialectología, como los atlas y las monografías dialectales.

El objetivo principal de este trabajo, como indica su título, es dar a conocer la contribución pionera del COSER al estudio de la variación gramatical del español y actualizar los informes previos presentados sobre este corpus rural. Su estructura es la siguiente: tras esta breve introducción (§ 1), se lleva a cabo una descripción general del corpus, así como de la metodología empleada en su elaboración (§ 2). Después, se da a conocer el estado del COSER en línea, a día de hoy (§ 3). Posteriormente, en el apartado principal del estudio, se resumen y sintetizan las aportaciones concretas que se han realizado hasta la fecha sobre fenómenos gramaticales de variación del español gracias a los datos del COSER (§ 4). Las conclusiones finales cierran el trabajo (§ 5).

2. DESCRIPCIÓN GENERAL DEL COSER Y METODOLOGÍA

Varias de las cuestiones generales y metodológicas sobre el COSER han sido previamente tratadas en otros trabajos (*cf.* Fernández-Ordóñez 2011, 2010a, 2010b, 2009a, 2007c y *corpusrural.es*), por lo que nos limitaremos en esta ocasión a ofrecer únicamente la información más relevante para su descripción general y la metodología empleada.

Lo primero que hay que recordar es que el COSER es un corpus de entrevistas semidirigidas restringido al habla rural de informantes mayores de escasa escolarización, que han nacido y que todavía viven en la ruralia, los pueblos de España, es decir, está dedicado a las personas que fueron objeto de interés en la dialectología tradicional. Sin embargo, aunque el COSER se nutre del mismo tipo de informantes que los atlas lingüísticos y que muchas de las monografías dialectales publicadas, la metodología empleada y los objetivos son muy diferentes. La diferencia más importante es, quizá, que el desarrollo de la entrevista semidirigida grabada permite investigar el empleo de cualquier fenómeno gramatical en su contexto de uso. En efecto, en lugar de las oraciones aisladas y descontextualizadas propias de un cuestionario lingüístico, la entrevista recoge oraciones emitidas dentro de un discurso real, en el que se pueden investigar, además,

¹ Las conclusiones obtenidas del contraste entre grupos lingüísticos resultan metodológicamente adecuadas si tenemos en cuenta el tipo conversacional en que se obtuvieron los datos, tanto en el COSER como en otros corpus lingüísticos: la entrevista, tipo de conversación sometido al intercambio pregunta-respuesta.

los valores contrastivos, las motivaciones afectivas y las inferencias pragmáticas asociadas a una determinada estructura sintáctica².

Las distintas campañas de encuesta (desde 1990 hasta noviembre de 2018) han permitido obtener 1656 entrevistas en 1325 enclaves rurales de 52 provincias e islas de España, lo que suma un total de 1752 horas de grabación. Todo ello hace que la densidad de la red de puntos encuestados y de los datos recogidos sea muy superior a la del único atlas lingüístico nacional o de gran dominio (ALPI) y a la de los distintos atlas lingüísticos (y etnográficos) regionales (ALCyL, ALEA, ALEANR, ALECant, ALECMán y ALEICan). El número de informantes también es mucho mayor, pues el COSER cuenta a día de hoy (enero de 2019) con 2574 informantes registrados: 1351 mujeres (52,4 %) y 1223 hombres (47,5 %). La edad media es de 73 años, siendo ligeramente menor en las mujeres (72,2) que en los hombres (73,7). El número de hablantes de cada sexo es estadísticamente representativo, por lo que permite investigar también diferencias lingüísticas asociadas a este factor.

La recogida de los datos se ha efectuado a ritmo anual desde el verano de 1990 (con una prueba piloto en León en 1988). En septiembre de 2012 y de 2013 se llevaron a cabo dos grandes campañas de encuesta para terminar el sur y el este peninsular, y desde 2016 se vienen realizando las encuestas en las islas Canarias y en Baleares. En total, se han organizado hasta el momento 48 campañas (*cf.* corpusrural.es/campanas.php).

Por lo que respecta a la metodología empleada, en el COSER se prioriza la entrevista oral semidirigida por parte del equipo de encuestadores, durante la cual uno de ellos suele guiar la conversación, primero hacia ciertos módulos temáticos de la vida rural tradicional como la matanza del cerdo y los cultivos propios del lugar. No obstante, pasado un tiempo y ganada la confianza del informante, la conversación se desplaza hacia otros temas como la vida en el pueblo, la educación recibida, las experiencias personales y la familia, dependiendo siempre del grado de comodidad y de espontaneidad que muestre el informante. La entrevista se centra en estos temas para facilitar que el informante potencial acepte ser entrevistado, al reconocer que está en posesión de ciertos conocimientos sobre un sistema de vida en decadencia, producto tanto de su experiencia personal como de su edad. Todo ello le confiere una autoridad informativa ante el entrevistador que viene de la capital. Los informantes aceptan la entrevista al comprender el interés por el testimonio de su sistema de vida, del que pocos guardan recuerdo y del que es experto.

Por lo general, el primer contacto con los informantes se realiza sin ninguna gestión previa, nada más llegar al enclave seleccionado entre los habitantes del lugar que reúnan los requisitos mencionados (persona mayor de escasa escolarización, que haya nacido y vivido en el pueblo y a la que le guste hablar). Todo investigador que haya realizado trabajo de campo sabe que el éxito de una entrevista no está asegurado, y que una entrevista puede ser mejor o peor bajo las mismas condiciones de partida. A este respecto cabe señalar que no todas las entrevistas son igualmente informativas, sino que su calidad depende de la disposición del informante, de la habilidad de los

² Las carencias de los atlas lingüísticos y de las monografías dialectales en lo concerniente al estudio de la gramática no deben, sin embargo, conducir a descartar su utilidad, ya que cuando algunas de estas obras se concibieron y realizaron ni la sintaxis ni la sociolingüística habían alcanzado el desarrollo teórico que han experimentado en los últimos cincuenta años. Aunque la metodología del cuestionario geolingüístico empleada para reflejar el habla en el ALPI y los atlas regionales posteriores es muy diferente de la metodología de la entrevista sociolingüística del COSER, debe reconocerse que ambas son hijas del estado teórico de la dialectología de su tiempo. El desarrollo de la sociolingüística ha puesto de manifiesto las múltiples limitaciones de la metodología de los atlas lingüísticos, pero hay que tener en cuenta también que, puesto que no existen grabaciones del habla de épocas pasadas equivalentes a las actuales, sus datos siguen constituyendo un testimonio válido, por imperfecto que sea, para estudiar la variación, incluida la gramática (*cf.* Pato 2003, Heap 2006, entre otros trabajos).

entrevistadores y de la interacción entre ambos. En algunos casos extremos solo es posible identificar *a posteriori* a los informantes que, por circunstancias varias –como haber pasado largas temporadas fuera del pueblo o haber disfrutado de una formación escolar mayor que el resto–, no encajan lingüísticamente ni en su área ni en su perfil sociolingüístico.

Respecto al número de informantes encuestado en cada enclave, para el COSER se suele entrevistar a un solo hablante en profundidad, ya sea hombre o mujer. Esto no impide que, en ocasiones, la entrevista pueda incluir esporádicamente a otros individuos, generalmente miembros de la familia y conocidos del entrevistado, que, atraídos por la circunstancia extraordinaria de la entrevista, se sienten tentados a intervenir y dar también su testimonio.

La duración media de las grabaciones realizadas es de aproximadamente una hora y tres minutos por enclave. Sin embargo, algunas de ellas cuentan con media hora y otras con hasta más de dos horas y media. Como todo investigador de campo sabe, la calidad de los datos no es directamente proporcional a la duración de una entrevista, ya que existen grabaciones excelentes y altamente informativas de media hora, cuyos resultados son equiparables a los obtenidos en una entrevista de mayor duración, y viceversa.

En cuanto al procesamiento de los datos, en 2009 se digitalizaron todos los materiales anteriores que estaban en formato analógico (desde 1990). Desde 2012 las entrevistas se graban tanto en audio como en vídeo, siempre con el consentimiento oral del informante.

La edición de las transcripciones ha seguido los avances de la tecnología, desde la transcripción a mano o a máquina hasta la informatización total de las mismas, como parte de diversos proyectos de investigación y de los trabajos finales de curso para las asignaturas que, en sucesivos planes de estudio en la Universidad Autónoma de Madrid, han tratado la dialectología del español desde 1990 hasta 2018. Las primeras transcripciones, de carácter bastante libre, se convirtieron en transcripciones pautadas desde 2005, con indicación de distintas marcas de conversación, temas tratados y fenómenos dialectales documentados. Las transcripciones disponibles están ahora todas homogeneizadas y se ha empleado para ello un sistema de ortografía convencional en el que se reflejan algunas características propias de la oralidad, como la omisión de algunos segmentos fonológicos (*comprao* por *comprado*) y la adición de otros (*muncho* por *mucho*), así como algunas acentuaciones dialectales (*pájaro* por *pájaro*). Todo ello permite proporcionar una imagen algo más cercana del discurso de los informantes rurales y de su habla, pero manteniendo globalmente las convenciones generales de la ortografía y favoreciendo la legibilidad de los textos (*cfr.* <http://www.corpusrural.es/transcripcion.php>)³.

Se reseñan, además, con marcas específicas las hablas simultáneas, las conversaciones cruzadas, otros elementos como risas, toses, chasquidos, elementos de comunicación no verbal –como los gestos–, así como la mayor o menor inteligibilidad de lo transcrito. También se indican las interrupciones (que pueden servir al hablante de herramienta para planificar el discurso) y las autocorrecciones (como indicadores de conciencia sociolingüística), entre otros muchos aspectos (*cfr.* <http://www.corpusrural.es/marcas.php>).

La decisión de transcribir en ortografía convencional algunos aspectos de la pronunciación dialectal obligó a crear un sistema de desambiguación de algunas formas

³ Aspectos como la pérdida de [d] intervocálica, la pérdida de algunas consonantes finales (-r, -l, -n, -d), la asimilación de [r]l/ [rs] > [l]/ [s] en los infinitivos con clítico o el desplazamiento acentual se aceptan transgrediendo la ortografía convencional. No obstante, la mayor parte de los cambios fonéticos o fonológicos no se reflejan: seseo, ceceo, yeísmo, neutralización, glotalización y pérdida de -s, -z, etc.

para su posterior etiquetado morfosintáctico. Por ejemplo, la transcripción de la pérdida de la *-d-* intervocálica y de la *-r* final condujo a desambiguar los participios en *-ada* y los infinitivos de la primera conjugación *-ar* para que el lematizador no pudiera confundirlos⁴.

Editar el COSER ha implicado tomar decisiones importantes sobre ciertas convenciones necesarias. En efecto, ya que desafortunadamente no contamos con un sistema universalmente aceptado para todos los corpus orales, y que el proceso de transcribir es una tarea hermenéutica sujeta a errores de audición y de comprensión, la uniformidad deseable supone siempre un reto. A esa dificultad debe añadirse que la transcripción y revisión es una tarea colectiva en la que han participado muchos estudiantes y colaboradores a lo largo de los años.

Desde 2016 las labores de gestión del proyecto se realizan en una base de datos en línea, en la que, con distintos permisos de acceso, pueden editarse y transcribirse las grabaciones y someterlas a revisiones sucesivas de forma centralizada⁵.

3. EL COSER EN LÍNEA, HOY

Por lo que respecta a la difusión del corpus, desde 2005 a 2015 se alojó en la página web www.uam.es/coser. En 2015 migró a www.corpusrural.es. En esta página se incluye toda la información general sobre el corpus (descripción y mapa de localidades encuestadas), la recogida de los datos (metodología y campañas), la transcripción (marcas de la conversación, transcripción ortográfica, temas tratados), los recursos disponibles (archivos de audio y texto así como muestras dialectales breves de aspectos sujetos a variación) y todo lo relativo al proyecto del COSER (su historia, los participantes, la financiación, las publicaciones, cómo colaborar con el proyecto, y los agradecimientos). Además, se habilitó una consulta básica de las grabaciones entonces transcritas. Desde el año 2017 esa web incorporó una consulta avanzada en pruebas que se ha actualizado, con diversas mejoras, en enero de 2019. En 2018 se completó la información de cada uno de los 1325 enclaves con la ficha técnica relativa a las entrevistas realizadas (código, encuestadores, informantes, duración, fecha de la encuesta), con independencia de que estén o no disponibles como archivo de texto y audio y de que sean accesibles a las consultas básica y avanzada.

Desde 2019 se ofrecen en la web, en texto, audio e interrogables mediante la consulta, 175 entrevistas correspondientes a 169 enclaves, que suman más de 229 horas de grabación y 4 771 943 palabras de texto transcrito, un 13 % del total del corpus⁶. Componen una muestra representativa de las siguientes provincias: Valladolid (6 enclaves), Burgos, Cáceres, Cantabria, León, Palencia, Toledo, Vizcaya, Murcia (5 enclaves), Álava, Albacete, Alicante, Almería, Asturias, Ávila, Badajoz, Cádiz, Castellón, Ciudad Real, Córdoba, Cuenca, Guipúzcoa, Granada, Guadalajara, Huelva, Huesca, Jaén, La Rioja, Madrid, Málaga, Navarra, Salamanca, Segovia, Sevilla, Soria, Teruel, Valencia, Zamora, Zaragoza (4 enclaves) y Lérida (1 enclave)⁷.

⁴ Un ejemplo concreto de ello sería el caso de *amarrá*, por el adjetivo/participio *amarrada* y por el verbo en infinitivo *amarrar*, por lo que tendríamos dos formas posibles: *amarr(á=ada)* y *amarr(á=ar)*.

⁵ <https://corpusrural.fe.uam.es/coser-editor/login.php>, de uso interno para los colaboradores del proyecto. Esta base de datos permite asimismo controlar las estadísticas.

⁶ Aparte de esas 229 horas transcritas y revisadas (13 % del corpus), contamos con otras 322:45 horas transcritas, sin revisar, que corresponden a la transcripción completa de 344 entrevistas, y otras 177:58 horas transcritas, que corresponden a la transcripción parcial de 304 entrevistas (cuya duración suma 311:30 horas). En total, 450:43 horas pendientes de revisión, un 25,6 % del corpus, que, sumado al 13 % ya revisado, alcanza la cifra de 38,6 % del total.

⁷ Las transcripciones que ofrezcan una muestra equivalente de Barcelona, Gran Canaria, Fuerteventura, La Gomera, Lanzarote, Tenerife, Tarragona, Mallorca, La Coruña, La Palma, Lugo, Orense y Pontevedra están en proceso. De Gerona,

En la consulta básica el buscador permite interrogar los datos delimitándolos por «Opciones de búsqueda» (exacta, aproximada y por expresión regular) y ordenar los resultados por provincia, frecuencia por provincia y frecuencia por enclave. Entre los filtros de búsqueda están el contexto de la búsqueda (entrevistas completas, informantes, hombres o mujeres), la provincia y la década de nacimiento del informante, las marcas (asentimiento, pausa, silencio, conversación cruzada, habla simultánea, literatura popular, nombre propio, gesto, risa, pronunciación, onomatopeya, ruido, chasquido, otras emisiones) y los temas (por orden alfabético: agricultura; alimentación; animales domésticos; bodas y noviazgos; caza y pesca; construcción y casas; costumbres y tradiciones; economía; educación y escuela; ejército y servicio militar; familia; fiestas populares; ganadería; industria básica; labores del hogar; matanza del cerdo; oficios; sanidad y salud; vida religiosa). También es posible el cartografiado dinámico de los resultados obtenidos tras la consulta.

En la consulta avanzada (*cf.* Imagen 1) la búsqueda (opcionalmente sensible a las mayúsculas) puede realizarse por forma ortográfica, lema y etiqueta gramatical (de acuerdo con el etiquetado proporcionado por FreeLing, <http://nlp.lsi.upc.edu/freeling/node/1>). Los filtros de búsqueda de la consulta básica se mantienen igual. En cuanto al uso de FreeLing, para los efectos prácticos del COSER se ha debido modificar el tokenizador (con etiquetas XML), se ha ampliado considerablemente su diccionario léxico y se han incluido las anotaciones de sufijos, terminaciones verbales y clíticos de las variedades dialectales (*cf.* De Benito, Pueyo & Fernández-Ordóñez 2016 para los problemas de transcripción y lematización).

FIGURA 1. Consulta avanzada del COSER

Desde 2019 todos los resultados se pueden descargar en formato Excel, son imprimibles y el cartografiado se realiza de forma dinámica y automática según el interés y las necesidades del investigador (con posibilidad de eliminar del mapa los enclaves que no se correspondan con lo buscado). Las ventajas que ofrece el COSER son varias, ya que permite un acceso rápido y fácil a la información (además de gratuito) y la posibilidad de acceder a todo el soporte original (el archivo audio de cada entrevista, *cf.* <http://www.corpusrural.es/archivos.php>). En un futuro se espera completar la alineación de texto y audio de todas las entrevistas disponibles, la mejora de la

Ibiza, Menorca y El Hierro todavía (enero de 2019) no hay materiales recolectados.

lematización y el etiquetado actual, así como la implementación de un etiquetado sintáctico basado en constituyentes o en dependencias⁸. La muestra de entrevistas disponibles, que hoy cubre toda la península ibérica, con la salvedad de Galicia y Cataluña, seguirá aumentando con el objeto de tener representación de esas áreas, así como de las islas Baleares y Canarias.

4. APORTACIONES BASADAS EN EL COSER

Como quedó señalado en (§ 1), el COSER es un corpus creado para el estudio de la variación que pueda encontrarse en el habla de los grupos socioculturales de menor educación del ámbito rural. El hecho de que la mayor parte de los corpus orales del español se nutran de fuentes como los medios de comunicación proporciona cierta singularidad al COSER y hace su aportación especialmente útil, pues facilita el estudio de las soluciones gramaticales no estándares, las cuales suelen ser sistemáticamente evitadas en la lengua escrita y en el habla de los grupos socioculturales de mayor educación. No son pocos los aspectos de variación gramatical que han sido estudiados a partir de los datos del COSER. Nos centraremos en los siguientes.

4.1. Morfología verbal dialectal

La morfología de la flexión verbal es un aspecto altamente sujeto a la variación dialectal que, sin embargo, cuenta con pocos estudios geográficos y teóricos, quizá porque generalmente no penetra en la lengua estándar y que, por ello, suele ser recesiva y difícil de documentar. Los atlas lingüísticos permiten hacerse una idea parcial de esta variación (*cf.* Mondéjar (1970) a partir del ALEA, o Llorente (1965) y Buesa & Castañer (1994) sobre el ALEANR), amén de las monografías dialectales. Sin embargo, hay pocos estudios recientes dedicados a ella en castellano (frente al interés y publicaciones que suscita la morfología dialectal del verbo en catalán, por ejemplo). Los datos del COSER, comparados con otras fuentes, han permitido definir o confirmar las áreas de uso, atestiguar la vitalidad y explorar las raíces históricas de algunos aspectos.

4.1.1. Los pretéritos perfectos fuertes

La 3ª persona del plural con desinencia analógica *-n* en los pretéritos perfectos fuertes del tipo *dijon*, *puson*, *estuvon* (Pato 2004a, 2010b), propia del leonés centrooriental, el castellano occidental y el extremeño, se conserva hoy de manera restringida, especialmente en León, Zamora, Salamanca, Palencia y Cáceres, y en puntos aislados y limítrofes de Valladolid, Ávila y Badajoz. La comparación con los datos del ALPI y otras fuentes anteriores a la Guerra Civil permite observar que en la segunda mitad del siglo XX su uso ha retrocedido en esas últimas provincias, así como en Burgos y Segovia. En ninguna fuente se registran las formas analógicas al norte de la cordillera Cantábrica, sino en los territorios repoblados al sur de la misma, es decir, el área originaria del posterior reino de León desconoce estas formas de perfecto. Ello unido a la ausencia en la documentación medieval leonesa del siglo XIII de este tipo de perfectos induce a pensar que se trata de un aspecto potenciado en época posterior, quizá gracias a la nivelación lingüística que tuvo lugar como efecto de la repoblación. Las primeras documentaciones de *hizon* o *pudon* en la zona occidental proceden de la literatura pastoril de Lucas Fernández, de finales del siglo XV.

⁸ Para ello contamos con el proyecto, dirigido por Miriam Bouzuita, «A Respeaking and Collaborative Game-Based Approach to Building a Parsed Corpus of European Spanish Dialects», financiado por el FWO (Medium-scale research infrastructure project: I000418N).

4.1.2. Los gerundios

Otra variante de flexión verbal estudiada gracias a los datos del COSER han sido los gerundios “análogos” (del tipo *supiendo, haciendo, tuviendo*), formas que siguen vivas en algunas de las variedades peninsulares (Pato & O'Neill 2013). En el ALPI las formas *tuyendo* y *quiendo* se registran especialmente en el área oriental, de norte a sur, desde Cantabria hasta Alicante, pasando por Burgos, La Rioja, Soria, Guadalajara y Cuenca, así como en Álava, sur de Navarra, Huesca, Zaragoza y Teruel. En el área occidental, se localizan de manera mucho más esporádica en diversos enclaves del sureste de León, centro de Palencia, Salamanca y norte de Cáceres. El COSER corrobora los datos del ALPI al mostrar que el área de conservación de estos gerundios es fundamentalmente oriental (registrados en las provincias de Álava, Soria, Guadalajara, Zaragoza y Teruel, con ejemplos también de Alicante y Castellón), siempre en informantes mayores. No obstante, el COSER también registra ejemplos en Cantabria y Palencia. Pato & O'Neill (2013) explican estas formas, en las que el tema de perfecto se extiende al gerundio, por condicionamiento fonológico del segmento temático [‘je], común tanto al perfecto (y tiempos afines) como al gerundio, en lugar de a través de la analogía proporcional con el gerundio del verbo *venir* (*viniendo*), tal y como se venía explicando anteriormente.

4.1.3. La acentuación

La variación acentual en formas del pretérito imperfecto de indicativo, subjuntivo y condicional (como *cantabámos* por *cantábamos*) (Pato 2012a) es otro de los rasgos estudiados. Estas formas que mantienen el acento latino en la 1ª y 2ª persona del plural han sido documentadas en el COSER para el pretérito imperfecto de indicativo en algunos enclaves de Burgos, La Rioja, Aragón (Huesca, Zaragoza y Teruel), Navarra, Álava, Guipúzcoa, Soria, Albacete, Almería, Castellón y Valencia, con algún ejemplo en Madrid, Toledo y León. En realidad, estas formas con acentuación etimológica se comportan como las del tema de presente (presente de indicativo y de subjuntivo), tiempos que no tienen acento columnar y que no mantienen la misma categoría morfológica acentuada en la 1ª y 2ª persona del plural. La conservación de la acentuación originaria, especialmente viva en Navarra y Aragón, puede deberse a que el castellano aragonés tiende a cambiar la acentuación esdrújula hacia formas paroxítonas (*sabádo* por *sábado*).

4.1.4. El pretérito imperfecto de indicativo

Además, los imperfectos de la 2ª y 3ª conjugación que mantienen o extienden análogicamente *-ba-* en su formación (*caíba, traíba*) también se documentan en el COSER en el área oriental del castellano (Pato 2018b). Tanto en la Península, Canarias como en América son más frecuentes con verbos cuya raíz finaliza en vocal y pueden por ello entenderse como una estrategia análogica, basada en el morfema *-ba* de la 1ª conjugación y apoyada por el imperfecto del verbo *ir* (*iba*), con la que se persigue deshacer el hiato originario. Su aparición independiente, en áreas no necesariamente vinculadas desde un punto de vista histórico, apunta a considerar este rasgo como un universal vernáculo (Chambers 2004).

4.1.5. La desinencia verbal *-éis* > *-ís*

Por último, la evolución de las desinencias de 2ª persona plural en *-éis* > *-ís* que documentan los atlas lingüísticos desde los años 30 del siglo XX se atestigua también en

los datos del COSER en toda el área del castellano oriental (Del Barrio 2018) y, además, ha hecho posible su cuantificación. De acuerdo con ella el foco de esta innovación o de su pervivencia parece localizarse en la provincia de Zaragoza, desde donde se extiende a las vecinas de Huesca, Teruel, Guadalajara, Soria y sur de Navarra, con presencia más reducida en Cuenca, Albacete, Ciudad Real, Toledo, Jaén y Almería. El análisis de los datos permite constatar que la reducción se da tanto en los verbos auxiliares (*habís*) y modales (*querís, podís*) como en los léxicos (*sabís, hacís, vis*). Asimismo, el tiempo verbal en que tuvo lugar la reducción primero es el presente de indicativo, común a todas las áreas. Desde ese paradigma la desinencia se extendió al presente de subjuntivo (*cantís*) y al futuro de indicativo (*cantarís*), extensión exclusiva de las zonas de mayor frecuencia.

Aunque todas estas formas de la morfología del verbo son recesivas en el español peninsular, los datos del COSER dan fe de su pervivencia en el ámbito rural a finales del siglo XX y principios del siglo XXI.

4.2. Concordancia de caso en los clíticos de objeto

4.2.1. *Leísmo, laísmo y loísmo*

La escasa tolerancia que la lengua estándar muestra hacia la variación gramatical hace que numerosos fenómenos dialectales se hayan visto sometidos con cierta frecuencia a un filtrado sociolingüístico, hecho que puede distorsionar los principios lingüísticos que explican su funcionamiento originario. Este ha sido el caso de los usos “anómalos” de los pronombres átonos, esto es del leísmo, el laísmo y el loísmo. Gracias a las entrevistas sociolingüísticas de Klein-Andreu (2000) y los datos del COSER (*cf.* Fernández-Ordóñez 1994, 1999) ha podido demostrarse que lo que los gramáticos percibían como usos desviados del empleo pronominal general son, en realidad, muestras parciales de paradigmas pronominales alternativos que se manifiestan plenamente en el habla de los grupos socioculturales bajos y en donde la selección del pronombre se realiza de acuerdo con principios lingüísticos diversos a los vigentes para el español estándar. Es decir, las hipótesis explicativas se habían creado únicamente sobre los datos que ofrecía la lengua escrita culta, hecho que dificultaba su correcta interpretación.

Si bien es cierto que en las descripciones previas se había percibido la conexión entre el leísmo, el laísmo y el loísmo (*cf.*, ente otros, Lapesa 1968), estos usos no se sometían fácilmente a una única explicación. Por un lado, la tendencia a distinguir los objetos directos personales (con *le* y leísmo personal) de los no personales (con *lo, la*, sin leísmo) explicaba el leísmo personal, pero dejaba sin aclarar los motivos por los que esta confusión afecta fundamentalmente a los objetos masculinos ni las razones por las que puede acompañarse de leísmo no personal (o de cosa), laísmo y loísmo. Por otro lado, la tendencia a asignar los pronombres recurriendo al género de su antecedente, sin considerar su función sintáctica, explicaba el leísmo masculino (personal y no personal) y el laísmo, pero no por qué el leísmo no acababa de establecerse plenamente para todo tipo de objetos masculinos, siendo siempre más frecuente cuando el antecedente era un objeto personal. Tampoco explicaba que el leísmo no se estableciera plenamente en el plural, donde además contendía con el loísmo.

A este respecto, la aportación del COSER ha sido fundamental para aclarar este aspecto mal comprendido de la sintaxis del español (*cf.* Fernández-Ordóñez 1993, 1994 y 1999, Fernández-Ordóñez 2001 desde el punto de vista histórico, y Fernández-Ordóñez 2012 para la gestación de las áreas dialectales). Se han acotado geográficamente las áreas que presentan leísmo, laísmo y loísmo, al tiempo que se ha demostrado que la aparente

falta de coherencia en los usos documentados responde a la existencia de varios paradigmas pronominales dialectales, alternativos al general del español⁹, empleos que se mezclaban en los estudios previos distorsionando su interpretación¹⁰. Con todo, además de soluciones de transición, existen tres grandes paradigmas (de los que solo se reseñan abajo los empleos discrepantes del general):

- a) El del romance hablado en contacto con el vasco, caracterizado por presentar leísmo asociado a objetos animados, masculinos y femeninos (*A Jon/ A Miren le veo*), y pronombres nulos para los objetos inanimados (*La bicicleta Ø tengo ahí*).
- b) El cántabro, caracterizado por presentar leísmo con los objetos contables masculinos en singular (*A Juan le veo; El libro le tengo*) y reservar *lo* para los objetos no contables, masculinos y femeninos (*El vino/ La leche lo bebo*).
- c) El castellano, también llamado paradigma referencial, caracterizado por haber eliminado las distinciones de caso entre objetos directos e indirectos. Los pronombres se asignan de acuerdo a su género, número e interpretación \pm contable. Así, hay leísmo para contables masculinos en singular (*A Juan le veo; El libro le tengo*) y, en la mitad norte del área, también en plural (*Los niños les veo; Los libros les tengo*); *lo* para objetos no contables, masculinos y femeninos (*El vino/ La leche lo bebo*), también cuando son indirectos (*A la leche lo echan de todo*); laísmo para los objetos indirectos femeninos (*La doy un libro; Las gusta el cine*); y loísmo para los objetos indirectos plurales en la mitad sur del área (*A los niños los gusta jugar*).

Mientras que el paradigma vasco representa la materialización de la hipótesis del leísmo como extensión del dativo a los objetos personales, tanto masculinos como femeninos, los paradigmas cántabro y castellano se caracterizan por basar la selección pronominal en la categorización semántica del antecedente como contable o no contable, categoría lingüística que no se había considerado anteriormente y que explica el hecho de que el leísmo fuera universal con los antecedentes personales masculinos (siempre contables y referidos por *le*) pero no se generalizara con los no personales, ya que estos pueden ser contables (referidos por *le*) o no contables (referidos por *lo*). El paradigma castellano, a su vez, se distingue del cántabro por haber eliminado la categoría de caso, generalizando *la(s)* y *lo* como pronombres de dativo. Además, en el plural masculino presenta dos soluciones diferentes según el pronombre preferido: a) *les* al norte (en el noroccidente de Burgos, Palencia y Valladolid); y b) *los* al sur (en el este de Salamanca y Cáceres, Ávila, oeste de Toledo y norte de Madrid). Los territorios en que se emplea este

⁹ El paradigma del español estándar (y el hablado en la mayor parte de variedades americanas) emplea *lo(s)* para OODD masculinos (*A Juan lo veo; El Libro lo tengo*) y neutro (*Eso no lo sé, la(s)*) para OODD femeninos (*A María la veo; La caja la tengo*) y *le(s)* para OOI (*A María / A Juan le regalé un libro; A eso no le doy importancia*).

¹⁰ Los diferentes tipos de leísmo tradicionalmente descritos son: a) OD masculino de persona singular (*¿Conoces a Juan? Sí, le conozco hace tiempo*); b) masculino de cosa singular (*¿Sabes dónde está mi libro? No, no le he visto por aquí*); c) de persona plural (*Esta tarde voy a recoger a los niños del colegio y les llevaré al parque*) y de cosa plural (*Fui a buscar los discos que quería y les encontré en la tienda de abajo*); d) femenino de persona singular (*A María hace tiempo que no le veo*) y plural (*Aquí no hay monjas. En la guerra les mataron a todas*). El laísmo de persona puede ser singular y plural (*Cuando vi a Pepa, la di su regalo; A las niñas de hoy ya no las gusta coser*) y de cosa (*Coges la sartén, la das la vuelta y ya tienes lista la tortilla*). Y el loísmo de persona plural (*Cuando recojo a los niños del colegio, los llevo la merienda*) y de cosa singular y plural (*Cuando el arroz está cocido, lo echas la sal; Para arreglar esos trajes, hay que sacarlos el bajo*), y de objeto neutro (*Yo no lo doy ninguna importancia a eso*).

sistema son, pues, los del centro y occidente de Castilla, desde el sur de la cordillera cantábrica hasta La Mancha.

Esta perspectiva general ha sido algo matizada en un análisis detallado de los pronombres átonos en la provincia de Álava (Camus Bergareche y Gómez Seibane 2015). Los datos del COSER han demostrado que el sistema propiamente vasco se da sobre todo al norte de la provincia, en contacto con el euskera, mientras que el área occidental se aproxima más al uso castellano y la oriental al uso navarro-riojano.

4.2.2. El doblado de objetos directos

Otro de los fenómenos relacionados con los pronombres átonos ha sido el estudio de la duplicación de objetos directos posverbiales en el español en contacto con el vasco (*Le veo a Juan*) (Gómez Seibane 2017), teniendo en cuenta para ello variables sintácticas, semánticas y cognitivas. Sobre los datos del COSER de zona vasca, con un corpus de control del castellano central, se ha podido probar que el doblado de objetos directos posverbiales es mucho más frecuente en la primera área que en la segunda (71, 2 % vs. 51,5 %). Además, el tipo de objeto doblado por el clítico es mayoritariamente pronominal (52,9 %) e inanimado (67,6 %) en zona castellana (*Lo teníamos que hacer todo*) mientras que es nominal (76,1 %) y humano (64,3 %) en zona vasca (*Tenía que ir a ayudarle al marido*). La mayor extensión del doblado al norte se revela también en que los antecedentes del clítico son, desde un punto de vista cognitivo, objetos semiactivos e inactivos, mientras que los del corpus de control son mayoritariamente activos y semiactivos¹¹.

4.2.3. La pronominalización con el verbo haber

Por último, los datos del COSER han sido la fuente para estudiar la posible pronominalización del argumento que acompaña al verbo *haber* existencial, siempre de carácter indefinido e inespecífico (*Antes había médico pero ahora no lo hay*) (De Benito 2016). En primer lugar, se ha precisado el área geográfica que acepta el pronombre, no restringida al noroeste peninsular, tal y como se creía, sino que se extiende desde Asturias y Cantabria por todo el occidente y centro de la Península hacia el sur, dejando fuera la Castilla oriental y el este peninsular. Mientras que en las variedades centrooccidentales se acepta el pronombre en el 43,4 % de los casos, en las orientales ello solo sucede en el 6,7 %. Dos de los factores que influyen en esta pronominalización son la animación del argumento y el tipo de construcción existencial. De hecho, los objetos animados aparecen referidos por un clítico con mayor frecuencia que los inanimados (40,3 % vs. 31,7 %) y la aparición de los pronombres átonos es más común en las construcciones adscriptivas (en las que el argumento se acompaña de una predicación secundaria) (*Sillas las hay pequeñas y grandes*) (50 %) que en las de locación (*Sillas las hay*) (26,5 %). Por otro lado, de gran importancia teórica es la documentación esporádica de ejemplos de *haber* concordado en plural con el argumento pronominalizado (*También las habían otras*), que vienen a sostener la idea de que el argumento único de *haber* presenta características mixtas de objeto y de sujeto (*cf. infra*, 4.4.1).

¹¹ Los activos han sido mencionados de forma inmediata en la interacción, los semiactivos han sido mencionados con cierta distancia, pero son identificables a través de alguna vinculación textual o inferencial, y los inactivos son completamente nuevos y no son accesibles textual e inferencialmente.

4.3. Concordancia de género

4.3.1. El neutro de materia

En algunos de los casos que hemos revisado anteriormente los datos del COSER han hecho posible estudiar y entender mejor fenómenos dialectales que ya eran conocidos. El interés del COSER resulta mayor si atendemos al hecho de que ha permitido registrar fenómenos completamente ignorados en las gramáticas y monografías dialectales hasta ahora, como el *neutro de materia*, fenómeno que se suponía exclusivo del asturiano centrooriental y de algunos puntos de Cantabria. Sin embargo, esta concordancia se registra también en todo el centro y occidente de Castilla hasta Toledo (Fernández-Ordóñez 2006, 2007a, 2007b, 2009b, 2019).

Por *neutro de materia* se entiende tradicionalmente la expresión morfológica, en los pronombres y adjetivos concordantes, de la categorización del nombre como discontinuo (contable) o como continuo (no contable)¹². La concordancia está restringida a pronombres y adjetivos, y no se da en el nombre, los determinantes, cuantificadores o adjetivos pronominales (*el/ este/ mucho/ buen pelo; la/ esta/ mucha/ buena lana*, pero **lo/ esto/ bueno/ mucho pelo/ lana*, a pesar del carácter no contable del nombre). Cuando el nombre masculino o femenino se interpreta como no contable, los pronombres personales y demostrativos, el artículo con núcleo elíptico y los adjetivos que siguen al nombre, sean modificadores adyacentes o de carácter predicativo, emplean una morfología diversa que los masculinos y femeninos contables. Por ejemplo, *La lana blanco está sucio vs. La manta blanca está sucia; Lo blanco está sucio vs. La blanca está sucia; La lana blanco lo lavo vs. La manta blanca la lavo*¹³.

En los adjetivos, la expresión de la concordancia de materia presenta menor vitalidad que en los pronombres. Desde el punto de vista sintáctico, está restringida a los adjetivos postnominales y a aquellos situados en posiciones predicativas. La variable distribución geográfica y la cuantificación de la concordancia en los adjetivos referidos a nombres no contables femeninos ha permitido concluir que se manifiestan con intensidad diversa según la clase de palabra y la posición sintáctica: pronombre clítico de OD (81,4 %) > predicativo (59,2 %) > atributo con *estar* (54,9 %) > atributo con *ser* (31,3 %) > adjetivo modificador (13 %) ¹⁴.

Así, la concordancia es regular en los pronombres de objeto y bastante elevada en los predicados secundarios (*Esa lana de las ovejas se lava bien lavao*) y en los atributos con la cópula *estar* (*Cuando estaba la leña bien quemadito ya, pues lo dejabas, lo esparcías bien por todo el horno*). Estos dos tipos de predicados tienen en común que suelen denotar estados acotados o propiedades de estadio o episódicas. En cambio, la concordancia neutra descende drásticamente con los atributos introducidos por la cópula *ser* (*La leche de cabra es muy bueno, mejor que lo de vaca*) y con los adjetivos que modifican directamente al nombre (*Si lo coge el agua, si lo siegas y te lo coge el agua, ye una hierba blanquino que...*), quizá porque suelen denotar propiedades individuales.

¹² Este conjunto de concordancias no debe confundirse con la existencia de un tercer género neutro, ya que todos los nombres que reciben interpretación continua son, desde el punto de vista léxico, masculinos o femeninos. Se trata de una concordancia basada en rasgos de carácter semántico, no léxico, y equiparable a otras concordancias *ad sensum*.

¹³ Las únicas diferencias morfológicas en el área son las siguientes: en los clíticos de OD se emplea *lu* en Asturias y *le* en Cantabria y Castilla para los contables masculinos; en los adjetivos masculinos los contables acaban en *-u* y los no contables en *-o* en Asturias (*el coche rotu vs. el vino frío*) mientras que en Cantabria y Castilla todos terminan en *-o* (*el coche roto, el vino frío*).

¹⁴ Únicamente en Asturias la concordancia ha llegado a afectar al adjetivo postnominal y en alguna de las variedades asturianas puede manifestarse, si bien de forma irregular, en la morfología del nombre (Lena, Quirós).

La concordancia de materia propia de los dialectos peninsulares refleja el patrón de la jerarquía universal de la concordancia propuesta por Corbett (2006): *modificador < predicado < pronombre relativo < pronombre personal*. De hecho, el examen detallado de los datos dialectales de la concordancia de materia permite refinar la jerarquía propuesta, distinguiendo dentro de la posición *predicado* entre predicados de propiedades individuales y de propiedades episódicas, siendo este segundo grupo el que expresa antes la concordancia de base semántica: *adjetivo modificador < predicados individuales < predicados episódicos < relativo < pronombre*.

4.3.2. El sincretismo en los cuantificadores evaluativos

El COSER también ha servido para documentar formas de difícil registro y aumentar la casuística de los ejemplos. Es el caso del sincretismo femenino en los cuantificadores evaluativos del oriente cántabro, del tipo *mucha trabajo*, que suele aparecer con nombres de materia (*mucha vientu*) y con nombres abstractos e infinitivos (*poca talentu; tanta trabajar*) (Fernández-Ordóñez 2015). Esta extraña concordancia se documenta, según los datos del COSER, en un área muy reducida del este cántabro (Vega de Pas, Arredondo, Rubalcaba, Quintana de los Prados (Cantabria) y Orbaneja del Castillo (Burgos)). El sincretismo de género –y número– se registra en cuantificadores universales en no pocas variedades románicas y, aunque más raramente, también en los evaluativos, pero siempre a favor del masculino. La existencia de estos cuantificadores cántabros plantea la cuestión de si podrían representar restos del neutro plural latino como los que se documentan en napolitano antiguo (*tanta angustie* ‘tanta angustias’; *quanta homini* ‘cuanta hombres’).

4.4. Concordancia de número

Uno de los aspectos de variación gramatical más frecuentes en las lenguas es la variable concordancia de número vinculada al reanálisis sintáctico de algunos constituyentes. Los datos del COSER han hecho posible estudiar algunos ejemplos, ya conocidos, pero ni cuya distribución geográfica en la Península ni sus factores condicionantes estaban bien definidos.

4.4.1. La concordancia de haber existencial en 3ª persona

La construcción impersonal existencial del verbo *haber* con un único argumento como OD tiende a reanalizarse en español de forma que ese argumento se interpreta como sujeto e induce la concordancia plural en el verbo (*Había casas > Habían casas*). Esta concordancia es muy frecuente en América, Canarias y en el español en contacto con el catalán, mientras que en el resto de la península ibérica se considera propia del habla popular o poco prestigiosa, sin haber sido objeto de un estudio de conjunto. Los datos del COSER han permitido demostrar la amplia difusión geográfica peninsular de las formas pluralizadas, que había escapado, por su baja frecuencia, a los atlas (Pato 2016, Bouzouita & Pato 2019). Si bien la concordancia alcanza su mayor frecuencia en el área valenciano-catalana (65 %), también existe en Murcia (46 %) y, esporádicamente, en porcentajes inferiores al 3 %, en otras comunidades (Cantabria, Extremadura, Castilla-La Mancha, Andalucía, Aragón, Castilla y León y el País Vasco). El contacto lingüístico con el valenciano favorece las formas plurales. También es frecuente en las construcciones perifrásticas (*Pueden haber problemas*). Aunque se ha observado para América que el rasgo semántico [+ humano] del argumento propicia la concordancia (*Han habido chicos que...*), tal diferencia no es tan clara en el español europeo. La importancia de la semántica se observa en que la concordancia puede darse incluso con

relativos (*Habían quien iban*) o nombres colectivos (*Habían gente*) [+ humanos] de semántica plural y número singular.

4.4.2. La concordancia de haber existencial en 1ª persona plural

Muy relacionada con la concordancia en 3ª persona plural de las construcciones existenciales figura aquella en que el verbo *haber* concuerda en 1ª persona plural con su argumento con lectura inclusiva del hablante (*Los que habemos no cabemos en el mundo*) (Castillo Lluch & Octavio de Toledo 2016). Mientras que es muy frecuente en América, en España se daba como típica del habla popular, sin más precisiones. Los atlas lingüísticos apenas documentan esta concordancia, por lo que los datos del COSER han hecho posible reconstruir por vez primera su distribución geográfica peninsular y atestiguar su empleo en distintos tiempos (*habíamos, habremos, habríamos, hemos habido*). La concordancia aparece sobre todo en las áreas periféricas de la Castilla central, desde León a Andalucía occidental, pasando por Extremadura, y desde Huesca a Andalucía oriental y Murcia, sin olvidar las zonas castellanohablantes valencianas. El rasgo semántico [+ humano] del argumento es obligado en la construcción que comenzó a ser plenamente existencial solo una vez que *avemos* perdió del todo su empleo como verbo posesivo, desde mediados del siglo XVII. La distribución geográfica de *habemos* existencial parece tener que ver con la pervivencia, en los mismos territorios y hasta el siglo XIX, de *habemos* auxiliar, mientras que en Castilla ya en el siglo XVI *habemos* había sido reemplazado por *hemos*.

4.4.3. La concordancia en los infinitivos

Otro ejemplo de expresión variable de la concordancia de número lo proporcionan los infinitivos (o gerundios) de verbos o construcciones reflexivas de sujeto plural (*Tenían que irsen para | porque no estuvieran las chicas por los pisos metiéndose con hombres*) (Heap & Pato 2012, con datos no completos). Esta concordancia en las formas no personales de los verbos reflexivos es un rasgo propio de la zona oriental (navarro-aragonesa y castellana), según los datos del ALPI que corrobora el COSER; no es, por tanto, un rasgo general al español de las clases bajas, contra lo que se había creído. Aunque actualmente predominan las formas en *-se*, los datos del COSER han permitido demostrar que la concordancia afecta a todo tipo de verbos y construcciones reflexivas (cfr. *infra*, 4.5): los reflexivos directos, los únicos acreditados por el ALPI (*No quieren abrigarsen*), los reflexivos indirectos (*Hacersen unas mangas*), los recíprocos (*Casarsen no se casaron*), los anticausativos (*Los chorizos pa secarsen*), los conversivos (*Tienen que venir a preocuparsen un poco*) e incluso los verbos intransitivos (*Salirsen de casa*) y transitivos (*Empiezan a comprarsen cosas*) sin cambio valencial, además de los no reversibles (*No vinieron más que a pitorrearsen de nosotros*). Además, la concordancia se extiende a la pasiva refleja, aunque el argumento no sea específico (*Ya echaron a formarsen sindicatos*).

4.5. Construcciones reflexivas, medias e impersonales con se

Las grabaciones del COSER son la base de uno de los estudios más completos y detallados efectuados hasta el momento sobre las construcciones reflexivas, mediopasivas e impersonales con *se* en español actual (De Benito 2012, 2013, 2015a, 2015b). La novedad reside en que, por vez primera, hay una medición cuantitativa de la presencia o ausencia del reflexivo en los distintos tipos de construcciones que pueden aceptarla. La cuantificación permite hipotetizar sobre las vías de extensión del reflexivo, desde su locus original, las construcciones reflexivas, a los demás tipos y medir, además,

si esa extensión fue condicionada por parámetros sintácticos o más bien semánticos (De Benito 2015b).

4.5.1. Las construcciones reflexivas y recíprocas

El estudio ha permitido determinar el comportamiento sintáctico diferenciado de las construcciones reflexivas y recíprocas en atención a su contenido semántico. La presencia del reflexivo y de su refuerzo enfático *a sí mismo*, aunque posible en toda construcción definida como reflexiva, no es idéntica en términos porcentuales. El refuerzo se da más con aquellos verbos que pueden alternar entre una construcción reflexiva o no (*Juan se mira vs. Juan mira la tele*) o reflexivos ‘puros’. En cambio, los reflexivos ‘naturales’, aquellos en que la acción se dirige prototípicamente a uno mismo, como los de cuidado personal, aceptan con mucha menor frecuencia el refuerzo reflexivo *a sí mismo* (*Juan se peina / se lava / se ducha*). Esta diferencia semántica con consecuencias sintácticas tiene su reflejo en la distribución geográfica: en el área noroccidental el reflexivo puede faltar en los reflexivos naturales (*Juan peina*) pero siempre está presente en los puros (*Juan se mira*).

También la cuantificación del uso ha permitido captar sutiles diferencias sintácticas entre los distintos papeles semánticos del reflexivo indirecto. El reflexivo indirecto puede tener distintos roles: destinatario (*Me envié el paquete*), benefactivo (*Me he conseguido una entrada*) y posesivo (*Me he roto la pierna; Me pongo el abrigo en los hombros*). Mientras que el destinatario puede ser el sujeto u otro argumento (*Le envié el paquete*), los reflexivos posesivos indican prototípicamente que el OD o el complemento locativo es parte del sujeto. Por ello, es más fácil su ausencia, tal como sucede en las variedades noroccidentales (*Rompí la pierna; Pongo el sombrero*). En cambio, los reflexivos destinatarios no solo son comunes a todas las áreas sino que son los que aceptan más el refuerzo enfático *a sí mismo* (*Me envié el paquete a mí mismo vs. ?Me he roto la pierna a mí mismo*).

Las construcciones recíprocas también son sensibles a la interpretación semántica del verbo. Los verbos naturalmente recíprocos son generalmente simétricos (*Juan y María se encontraron*) y suelen presentar con menor frecuencia el refuerzo *uno al / con el otro, entre ellos / sí*, en contraste con lo que ocurre con los recíprocos ‘puros’, generalmente no simétricos (*Juan y María se miran*), que, por ello, presentan en mayor porcentaje el refuerzo (*Juan y María se miran el uno al otro*). Otra diferencia sintáctica entre estos dos tipos de verbos recíprocos es el hecho de que la construcción llamada “discontinua” (que presenta el sujeto en singular y codifica el otro miembro de la acción recíproca con un sintagma preposicional: *Juan se encontró con María*) es propia de los verbos naturalmente recíprocos. Tal como sucedía con los reflexivos indirectos, los recíprocos indirectos, generalmente dativos posesivos, son mucho menos frecuentes en el área noroccidental (*Juan y María dieron la mano*).

4.5.2. Las construcciones ‘medias’

La puesta a prueba con un análisis estadístico de las hipótesis manejadas para explicar los distintos tipos de *se* ha arrojado importantes conclusiones. Hay acuerdo en que el reflexivo sigue un camino de reanálisis sintáctico desde el punto de vista tipológico que obedece a este patrón: *reflexivo > recíproco > medio > pasivo > impersonal*. Sin embargo, dentro de los usos medios caben muchas configuraciones sintácticas y es lícito preguntarse por los mecanismos precisos y construcciones concretas en que tuvo lugar esa reinterpretación. Los usos llamados ‘medios’ se documentan en los siguientes tipos de alternancias diatéticas:

- a) La diátesis anticausativa, en que se elimina el agente y el paciente se promueve a la posición de sujeto (*Juan rompe la mesa* > *La mesa se rompe*).
- b) La diátesis conversiva, en que la causa expresada por el sujeto se degrada a complemento de régimen y el experimentante objeto se promueve a sujeto (*La noticia alegró a María* > *María se alegró de la noticia*).
- c) La diátesis deobjetiva, que implica la eliminación del objeto, generalmente abstracto y que adquiere una lectura genérica (*Juan explicó (algo)* > *Juan se explicó*).
- d) La diátesis antipasiva, que consiste en la supresión del objeto, el cual, a veces, puede expresarse como un complemento de régimen (*Juan aprovechó tu experiencia* > *Juan se aprovechó (de tu experiencia)*).

Además, el reflexivo se ha extendido a otros tipos de verbos que no manifiestan cambio valencial, tanto intransitivos (*ir* > *irse*) como transitivos (*comer algo* > *comerse algo*). Las diferencias porcentuales de expresión del reflexivo en cada tipo y en la geografía han permitido concluir que el camino de difusión fue el siguiente. La extensión del reflexivo comenzó con la diátesis conversiva, con uso casi categórico del reflexivo en todas las zonas y sujetos siempre animados (*María se asusta*), tal como las construcciones reflexivas. De ahí se extendió a la deobjetiva (*María se aguanta*) y a la diátesis anticausativa con sujeto animado (*María se levanta*). La difusión alcanza más tarde a la antipasiva (*Juan se olvidó*) y a la diátesis anticausativa cuando presenta un sujeto inanimado (*La ropa (se) seca*). Este itinerario se deduce de las frecuencias globales del empleo del reflexivo por tipo diatético: *conversiva* (99 %) > *deobjetiva* (87,2 %) > *anticausativa con sujeto animado* (80 %) > *antipasiva* (76,9 %) > *anticausativa con sujeto inanimado* (71,6 %).

Todas las configuraciones diatéticas con frecuencias más altas del reflexivo coinciden en la animación del sujeto, lo que las conecta con la diátesis originaria, la reflexiva. Por otro lado, el análisis de los datos permite configurar dos áreas dialectales: la noroccidental, en que la diátesis anticausativa presenta los índices de frecuencia más bajos, tanto para sujetos animados (56,33 %) como inanimados (42,3 %), y el resto del territorio, en que esos porcentajes se elevan notablemente (97,1 % y 75,5 %). Así, resulta evidente que el reflexivo es originariamente un procedimiento sintáctico intransitivizador del verbo que indica el cambio diatético.

4.5.3. El reflexivo sin cambio valencial

La difusión del reflexivo a los verbos sin cambio valencial es necesariamente posterior. A diferencia de la gran homogeneidad en el comportamiento sintáctico de los verbos anteriores, la dispersión léxica en el empleo del reflexivo es enorme. Incluso en los verbos que aceptan el reflexivo, su uso no supera el 25,1 % en los intransitivos (*salirse, irse, venirse, estarse*) y el 11 % en los transitivos (*comerse, saberse algo*). De nuevo, los sujetos animados favorecen la presencia del reflexivo (28 %) frente a los inanimados (15,8 %), al menos en los intransitivos, y el área noroccidental se resiste a aceptar estos usos más que el resto del territorio peninsular.

La extensión tardía del reflexivo a los verbos intransitivos sin cambio valencial puede explicarse como un proceso analógico basado en la semejanza de estos verbos con el contenido prototípico de los que experimentan esos cambios diatéticos (movimiento,

irse < *levantarse*, o emoción, *reírse* < *alegrarse*). Por lo general, el reflexivo indica un aumento de la agentividad o de la afectación del sujeto.

Necesariamente posterior debe haber sido su extensión a los verbos transitivos sin cambio valencial, en que el reflexivo parece dar cuerpo a un dativo benefactivo (*comerse algo*, *saberse algo*, *temerse algo*) que funciona como adjunto. Estos dativos también enfatizan la agentividad o la afectación del sujeto en el evento denotado. En realidad, apenas se distinguen del reflexivo que aparece en las construcciones reflexivas o recíprocas indirectas con un dativo benefactivo no argumental (*hacerse algo*), el cual, como sabemos, apenas aparece en el noroeste y que, por ello, se conforma como la fuente potencial de la analogía.

Además de reconstruir el camino por el que avanzó la reinterpretación del reflexivo mediante la cuantificación de su presencia en la estructura lingüística y en la geografía, otra importante conclusión teórica de este estudio es poner en tela de juicio la función supuestamente aspectual del reflexivo en los verbos sin cambio valencial. Para discutir esa interpretación, son pertinentes dos aspectos: por un lado, la ausencia de conexión diacrónica entre los empleos del reflexivo como indicador de un cambio diatético y su supuesta función de marcador aspectual de eventos culminados o logros; por otro, la gran cantidad de ejemplos en que el reflexivo no se relaciona con ningún cambio tal en la interpretación aspectual del evento (por ejemplo, *El agua se está saliendo*; *Me estoy terminando el café*; *Ayer me leía el periódico en un bar*).

4.5.4. Áreas dialectales vinculadas a los verbos reflexivos

Desde el punto de vista areal, este estudio identifica por vez primera una zona noroccidental en que los empleos del reflexivo están menos extendidos en general y, en concreto, en la diátesis anticausativa, en los reflexivos indirectos posesivos y benefactivos así como en los verbos intransitivos sin cambio valencial. Además, también se descubre y describe un área oriental, desde Castellón hasta Murcia, en que tiene lugar un sincretismo en el paradigma morfológico de los reflexivos en el plural a favor de *se* (*se casamos*, *se vais*) (De Benito, 2015a). Los datos muestran que el sincretismo se produce de forma más regular y geográficamente extendida en la 2ª persona de plural (*se queréis*) y que se ve favorecido por el uso del infinitivo en lugar del imperativo plural (*¡Sentarse de una vez!*). Aunque el fenómeno se da también en catalán, incluido el dialecto valenciano, las causas del cambio en esa lengua parecen fonéticas. En el castellano el contacto puede haber sido una motivación concomitante, pero también se ajusta al sincretismo derivado del empleo del infinitivo, extendido por toda Andalucía y claramente independiente del contacto bilingüe con el catalán. A estas áreas vinculadas a los reflexivos debe sumarse la oriental que expresa la concordancia plural del sujeto de infinitivos y gerundios mediante el reflexivo *-sen* (cfr. *supra*, 4.4.3).

4.5.5. Las construcciones impersonales reflejas

El *se* impersonal parece el último desarrollo de la extensión del reflexivo (*Aquí se vive bien*). Tradicionalmente se estima como prueba de la configuración impersonal su empleo con verbos intransitivos o, si el verbo es transitivo, que acepte la pronominalización del paciente siempre que sea humano y definido, esto es, vinculada al MDO (Marcado diferencial de Objeto) (*Se ve a Juan* > *Se le ve*). Ciertos dialectos aceptan, sin embargo, una pronominalización ampliada (De Benito 2012, 2013) en la que es posible referir con clítico a pacientes no humanos. Sobre los datos del COSER se demuestra que esa pronominalización se da exclusivamente en aquellas áreas en que los clíticos de 3ª persona tienden a eliminar las diferencias entre acusativo y dativo, la

referencial castellana, la cántabra y en el castellano hablado en contacto con el euskera. Por otro lado, gracias a la cuantificación de los datos se prueba que la pronominalización, nunca obligatoria, se extiende gradualmente siguiendo la Jerarquía de la animación: *humanos* (52, 5%) > *animados* (38, 8 %) > *inanimados contables* (12, 5 %) > *inanimados no contables* (6 %).

Otra importante aportación de este estudio ha sido descartar la hipótesis de que la pronominalización ampliada esté relacionada con un uso impersonal, esto es, no concordado del verbo con los objetos inanimados (*Se vende patatas ≠ Se las vende*). La desconexión entre los dos aspectos se prueba por la ausencia de coincidencia geográfica entre ambos. La falta de concordancia se da en áreas (como el Alto Aragón) en que la pronominalización ampliada es inexistente. Por otro lado, la topicalización del paciente definido no fuerza la pronominalización ampliada (solo el 4,2 % de los casos presentan ejemplos con copia pronominal como *Esa tela se la descose*), quizá porque en la mayoría de los casos se interpreta como sujeto de una pasiva refleja (*Esa tela se descose*).

El pronombre que figura en la impersonal refleja era originariamente un dativo, tanto para pacientes humanos masculinos y femeninos como para inanimados, y así continúa siendo en la mitad septentrional de América (*A María se le ve; La montaña se (le) ve*). En aquellas variedades peninsulares en que se han perdido total o parcialmente las distinciones de caso, se emplean, sin embargo, los pronombres de acuerdo con el paradigma pronominal en uso en cada variedad. En el País Vasco, solo aparece con *le*, y en el área referencial, los pronombres *le(s)*, *la(s)*, *lo(s)* seleccionados de acuerdo con los parámetros de género, número y ±contable que determinan su uso general en ese dialecto, si bien siempre pervive el *le* originario, compitiendo en minoría. Ello abre la posibilidad de que esté teniendo lugar una transitivización plena de la construcción en esos dialectos, aunque fuera originalmente intransitiva.

4.6. Variación modal

El COSER ha sido la base documental principal de estudios dedicados al desplazamiento del subjuntivo por el indicativo en la Península, en concreto, el empleo del condicional (o el imperfecto de indicativo) en lugar del pretérito imperfecto de subjuntivo en cualquier tipo de oraciones subordinadas (*Las costillas y todas esas cosas se metían en ollas para que se conservarían / conservaban*) (Pato 2003, 2004b, 2012c). El contraste entre los datos de los atlas y los datos del COSER confirma, en términos generales, la coincidencia en el área geográfica en que tiene lugar el fenómeno, pero resalta la incapacidad de los atlas en reflejar la variante minoritaria, el imperfecto de indicativo, y en detectar la existencia del mismo desplazamiento en las correspondientes formas compuestas del verbo. La ventaja del COSER es el elevado número de casos obtenidos (3588), lo que permite detectar la existencia de variantes minoritarias y cuantificar los datos para matizar su distribución. Así se ha podido identificar el área focal de este empleo (en la Castilla Vieja, el sur de Cantabria, el occidente de Vizcaya, el norte de Palencia, la Navarra media y las provincias de Álava y La Rioja) y las áreas de transición (en el norte de Cantabria, el oriente de Vizcaya, Guipúzcoa, la Navarra septentrional y meridional, el sur de Palencia, el este de León y el norte de Valladolid, Segovia y Soria). En las áreas en que tiene lugar este desplazamiento el condicional es la forma preferida (96 %) frente al imperfecto de indicativo (4 %) y es mucho más frecuente en los tiempos simples (61, 9 %) que en los compuestos (21, 6 %).

Por otro lado, la cuantificación ha hecho posible aplicar pruebas estadísticas como la regresión logística para valorar la influencia simultánea de varias variables sobre la manifestación del fenómeno (*cf.* Pato 2012c para una visión de conjunto). La presencia del condicional en la prótasis de las oraciones condicionales (*Si tuviera / tendría dinero,*

lo compraría) se había reconocido siempre como un uso dialectal estereotipado del habla de los vascos. Pero, en realidad, en el castellano septentrional la prótasis condicional no constituye el contexto originario de este fenómeno, sino el más visible a los gramáticos y dialectólogos. Gracias al análisis estadístico, se ha podido demostrar que la extensión de las formas indicativas en *-ría* y *-ba* a costa de las formas subjuntivas en *-ra* y *-se* tuvo lugar primero en las oraciones completivas, extendiéndose luego a las relativas y adverbiales relacionadas (modo, lugar y tiempo), y más tarde a las condicionales, para alcanzar finalmente al resto de contextos sintácticos (Pato 2003, 2004b): *completivas* (72,1 %) > *relativas* (61,7 %) > *condicionales / finales* (57,5 %).

Sin duda alguna, el orden *completivas* > *relativas* > *condicionales* pasó desapercibido porque en las oraciones completivas y en las relativas es posible encontrar alternancia de modo en el español sin que, a veces, sean nítidas las diferencias en la interpretación. En las primeras, las completivas, la alternancia de contenido modal suele explicarse por la presuposición del valor de verdad de aquello expresado en la subordinada, esto es si se presupone cierto o se afirma (*María sabía que su novio vendría a visitarla*, selección del indicativo) o no (*María deseaba que su novio viniera a visitarla*, selección del subjuntivo). Sin embargo, hay contornos que en los dos modos alternan de forma sutil sin que se identifiquen claramente esas diferencias de interpretación (*María esperaba / no creía que su novio vendría / viniera a visitarla*) y que se perfilan como el locus originario de la extensión del indicativo a costa del subjuntivo.

En cuanto al segundo tipo oracional, las relativas y adverbiales relacionadas, sucede algo semejante. La alternancia de modos se explica por el carácter más o menos específico del antecedente. Si es específico (existente o identificable) se emplea el indicativo (*El hombre, que sabría aquel oficio, había fallecido*), y si es inespecífico (no se afirma su existencia), el subjuntivo (*No hubo nadie que supiera aquel oficio*). Pero de nuevo hay lecturas en que los dos modos alternan con significado muy similar (*El hombre que sabría / supiera aquel misterio había desaparecido*). En consonancia con estos hechos, en el castellano septentrional la presencia del indicativo en lugar del subjuntivo es tanto más frecuente cuanto más específico es el antecedente. Un criterio formal indirecto permite medir la especificidad: la escala de la definitud. Aunque la interpretación de la especificidad de los sintagmas nominales definidos e indefinidos es contextual, los datos revelan una mayor presencia del desplazamiento acorde con ella: *antecedentes definidos* (64,5 %) > *antecedentes indefinidos* (58,3 %) > *nadie* (0 %).

Desde un punto de vista tipológico, la pérdida del subjuntivo según la escala *completivas* > *relativas* > *condicionales / resto* probablemente sea de aplicación a las variedades del español en América y a otras lenguas romances, ya que es un cambio predecible tanto en lo relativo al modo perdido (subjuntivo) como a los tiempos afectados (pasado), pues los elementos marcados tienden a perder la expresión flexiva de diferencias que se mantienen en los no marcados.

4.7. Rección

El estudio de la variación en la expresión del complementante *de* que rige un infinitivo subordinado, fenómeno conocido como *deísmo*, ha avanzado asimismo a partir del análisis de los datos del COSER (*Me han hecho mis padres de correr*) (De Benito & Pato 2015). En primer lugar, porque sobre el análisis del COSER se ha ofrecido el primer mapa que cartografía la extensión geográfica del deísmo, ya que los atlas apenas consiguen documentar este fenómeno. El mapa revela que hay ejemplos aislados de deísmo en la mitad septentrional peninsular, pero la expresión del complementante es más intensa y constante en la mitad sur (con ejemplos regulares en Andalucía, Murcia,

Extremadura, Castilla-La Mancha, Comunidad Valenciana y Madrid). Además, el número de ejemplos aumenta según nos acercamos hacia el sur, por lo que se configura como un rasgo sintáctico característico del español meridional.

Además, se han podido caracterizar los tipos de verbos con los que aparece: a) verbos de afección (*Me costó de adaptarme; Me da pena de verla*), a veces con lectura genérica; b) verbos de percepción (*Le oía de decir*); c) verbos de manipulación (*De bebérmelo no me dejó*); d) verbos de volición e intención (*Están deseando de venir*), e incluso; e) verbos modales en perífrasis verbales (*Suele de pasar*).

La categoría a la que pertenece de, preposición o complementante, ha sido debatida tanto en el español como en otras lenguas que presentan infinitivos regidos por la preposición, como el catalán, el francés o el italiano (De Benito & Pato 2015). Después de constatar la distribución geográfica desigual y paulatina de los tipos de verbos que aceptan el deísmo, estos autores deducen que los primeros en aceptarlo fueron los de afección, en los que el valor preposicional de *de* es perceptible en que a veces aceptan un complemento nominal (*Me da miedo del perro; Me da pena de mamá*), y los más antiguos documentados. A partir de esos empleos parece haberse gramaticalizado como complementante, ya que no introduce nombres (*Estoy deseando del regalo; *Le oía de sus pasos). En una etapa posterior, se habría convertido en partícula introductora de infinitivos, similar a *to* en inglés, fase solo presente en los verbos modales, mucho menos frecuente y más restringida en la geografía. Por tanto, esta hipótesis predice que hay diferentes tipos de deísmo según las áreas y las fases de evolución.

4.8. Fonética y gramática

La aparición de variantes fonéticas en palabras gramaticales o la incidencia de factores gramaticales en la evolución fonética también ha sido objeto de investigación con los datos del COSER.

4.8.1. El cuantificador *mucho*

El contraste entre los datos del ALPI y el COSER ha permitido conocer la extensión geográfica de la forma *mucho* (Pato 2013), fenómeno antiguo pero hoy de uso frecuente en Andalucía centrooriental, Murcia, Alicante y Albacete (con ejemplos esporádicos aragoneses y en las zonas castellanohablantes de Castellón y Valencia). Un área independiente se configura en Asturias. Esta *-n-* epentética se documenta también en otras lenguas y dialectos romances, como el asturiano y los dialectos septentrionales italianos, así como históricamente en francés. Documentado desde el español antiguo, su frecuencia de uso fue notable, sobre todo, en el siglo XVI.

4.8.2. La pérdida de *-d-* intervocálica

La pérdida de la *-d-* intervocálica, uno de los fenómenos más característicos de la fonética oral del español, ha sido estudiada en profundidad sobre los datos del ALPI y del COSER (Estrada 2017, 2018). Por primera vez, se ha podido analizar este cambio desde una perspectiva geográfica sobre mallas uniformes en la recolección de los datos. Como resultado, se ha delimitado el área geográfica afectada y cuantificado el alcance de la pérdida en distintos contextos fonotácticos, morfológicos y sintácticos.

La pérdida de la *-d-* es característica de la mitad sur de la Península. A partir del sistema central hacia el norte disminuye mucho, salvo en los participios en *-ado*, y solo se detecta con intensidad en un área aislada, la Asturias centrooriental y Cantabria, en la que influye el contacto con el asturiano. La isoglosa delimitadora es casi idéntica en el

ALPI y el COSER, lo que indica que, pese al tiempo transcurrido entre ambas muestras, se trata de una situación estable.

En cuanto a los factores que condicionan la pérdida, el estudio demuestra que: a) la pérdida apenas se da en contacto con una *glide* mientras que la apertura de las vocales en contacto con la *-d-* podría favorecerla (*a, o, e* vs. *i, u*); b) el acento previo a la *-d-* favorece su pérdida, así como la acentuación paroxítona; c) las palabras de tres o más sílabas también propician más la pérdida que las bisílabas; d) desde el punto de vista gramatical, la pérdida ordena su frecuencia de acuerdo a esta escala: *morfemas flexivos* (62,5 %) > *lexemas gramaticales* (37,9 %) > *morfemas derivativos* (30,3 %) > *lexemas* (19,7 %); e) es teóricamente importante constatar que los diminutivos eliminan la *-d-* siguiendo el modelo de su base (*delgado* > *delgao* > *delgaúcho*, pero no *delgaducho*); y f) la preposición *de* apenas se pierde, aunque desde el punto de vista fonotáctico se encuentre en las mismas condiciones intervocálicas que dentro de una palabra. Ello parece depender de si la preposición es parte de una locución lexicalizada (*ojo de gallo*; *lucero del alba*) o si enlaza dos nombres en un sintagma (*una caja de madera*), en cuyo caso la pérdida es casi inexistente. g) Finalmente, la frecuencia del elemento en que la *-d-* se incluye parece ser determinante. Por ese motivo, los nombres propios son especialmente renuentes a perderla mientras que los participios en *-ado*, los cuantificadores *todo* > *to, nada* > *na* o el verbo *poder* dan cuenta de la inmensa mayoría de los casos.

4.9. Léxico y gramática

En las entrevistas del COSER se registran adjetivos o adverbios poco o nada comunes en el habla común, tanto por conservación de formas arcaicas como por desarrollo de nuevos significados y posibilidades sintácticas. Algunos ejemplos estudiados son los siguientes.

4.9.1. Mucho como cuantificador de adjetivos y adverbios

La distribución geográfica del uso del superlativo absoluto con la forma plena *mucho* como cuantificador de adjetivos y adverbios (*mucho bien, mucho guapa*) ha sido también confeccionada gracias al COSER (Pato & Viejo Fernández 2017). Aparte de su empleo en Asturias occidental, es uso propio del español oriental del norte (con ejemplos desde el oriente de Burgos, La Rioja, Soria, Guadalajara, Navarra, Aragón –sobre todo Teruel– y Murcia). Como uso vigente en hablantes mayores de 70 años, es un fenómeno en retroceso y siempre en alternancia con *muuy*. No obstante, *mucho* parece conservarse mejor cuando modifica a un adverbio que a un adjetivo.

4.9.2. El adjetivo semejante

En el habla rural se documenta también un uso nuevo del adjetivo *semejante* (Pato 2018a). A los valores usuales de *semejante* como adjetivo calificativo ‘similar’ (*Un libro semejante al cuadro*), cuantificador ponderativo ‘tal’ (*No es lícito valerse de semejantes medios*) y determinante (*No he visto a semejante hombre*), se suma, en algunos enclaves de Madrid y Ciudad Real, un uso adverbial. Esta recategorización de adjetivo a adverbio parece derivar del segundo significado (‘de manera similar, de la misma manera, igualmente’). En los ejemplos registrados *semejante* modifica a adverbios deícticos de lugar (*aquí*) y modo (*así*) (*Tenías que tumbar la oveja, la oveja tripa arriba, el cuello aquí, la cabeza semejante aquí, y las patas pa abajo; El hombre o la mujer, cogía así un cacho, y lo metía semejante así, en el, en el, en el centro*). En estos ejemplos *semejante*

presenta un valor anafórico que hace referencia a una situación inmediatamente anterior, de modo que la paráfrasis de *semejante así* sería ‘de esta manera, por ejemplo’.

4.9.3. Los adverbios relativos e interrogativos

La distribución geográfica y sintáctica de los adverbios relativos e interrogativos *dónde*, *ónde*, *adónde*, *aónde* y *ánde* también ha sido estudiada sobre los datos del COSER y de los atlas lingüísticos (Del Barrio 2018). El ALPI permite descubrir una nueva isoglosa entre dos áreas dialectales ya conocidas por otros aspectos: la occidental, en que se emplea *onde* para la ubicación y *a(o)nde* para el destino, y la oriental, en que *a(o)nde* expresa ambos contenidos. El sincretismo en sentido contrario, a favor de *onde*, solo aparece en algunos puntos andaluces, extremeños y leoneses. Los datos del COSER revelan la desdialectalización acaecida a lo largo del siglo XX, pues *ónde*, *a(ó)nde* son ahora minoritarios en comparación con las formas estándar *dónde* y *adónde*. Aun así, *ánde* mantiene mayor vigencia en el área oriental y *ónde* en Asturias y Cantabria. Para expresar la ubicación *dónde* es la forma más frecuente (79,9 %) frente a *ánde*, *a(d)ónde* (20 %). Para la dirección, en cambio, *ánde*, *a(d)ónde* (58,8 %) superan a *(d)ónde* (43,1 %).

4.9.4. Otros cuantificadores y adverbios

Además, en las entrevistas del COSER se registran cuantificadores y adverbios supuestamente desaparecidos de la Península desde los siglos XVI y XVII. En Toledo aflora el cuantificador *algotro* (Octavio de Toledo 2016) y en varios enclaves de las provincias de Teruel y de Zaragoza aún está vivo el adverbio *agora* ‘ahora’, que ha perdurado sin interrupción en el ámbito aragonés (Pato 2010a).

5. CONCLUSIONES

La dialectología moderna sabe que la gramática dialectal es una fuente de información preciosa, hasta hace poco tiempo insuficientemente valorada, no solo para la caracterización de un dominio lingüístico concreto sino también para el estudio teórico y tipológico de las lenguas (*cf.*, entre otros, Chambers 1995, Kortmann 1999, 2004a, 2004b, Barriers, Corsnipp & van der Kleij 2002, Biberauer 2008). Como hemos visto a lo largo de este trabajo, los datos del COSER sirven para alcanzar interpretaciones más correctas de los principios lingüísticos operantes en las variedades orales.

El COSER constituye, además, un complemento tanto de los atlas lingüísticos como de los diversos corpus de habla culta y urbana que se han compilado, o se están compilando, en el mundo hispanohablante, con la posibilidad de estudiar cambios lingüísticos en tiempo real. Puesto que, como vimos, los informantes del COSER pertenecen al mismo grupo social de los informantes seleccionados para los atlas lingüísticos, resulta posible la comparación entre el habla de varias generaciones sucesivas de hablantes de las mismas áreas geográficas. Pese a las diferencias existentes entre una y otra metodología (cuestionario *vs.* entrevista oral), el contraste entre los datos permite investigar cambios en el tiempo transcurrido desde comienzos del siglo XX hasta comienzos del siglo XXI.

Esta metodología comparativa entre el COSER y los atlas y otros corpus para obtener conclusiones sobre la evolución diacrónica y la implantación sociolingüística de los usos sometidos a variación se ha puesto a prueba en distintos trabajos como los de Pato (2003, 2004b, 2010b, 2012c, 2013), Pato & Heap (2012), Pato & O'Neill (2013), Pato & Viejo (2017), Heap (2006), De Benito (2015b, 2016), Castillo Llach & Octavio de Toledo (2016), León Zurdo (2017), Estrada (2017), Lara (2018) o Del Barrio (2018), con

conclusiones relevantes sobre el mantenimiento o la desdialectalización de los aspectos estudiados (*cf.* §§ 4.1.1, 4.1.2, 4.1.5, 4.4.3, 4.6, 4.8.1, 4.8.2, 4.9.3).

De este modo, se ha podido constatar la desdialectalización de algunos aspectos propios de la gramática rural. Además de los citados *supra*, conviene mencionar el declive de las secuencias *me se* y *te se*, la de 2ª persona más frecuente que la de 1ª en extensión geográfica (Heap 2006), o del uso de *yo* y *tú* como pronombres regidos por preposición en el castellano de Aragón, rasgo común al aragonés, catalán y provenzal (Pato 2012b). Estas formas se documentan hoy día en el COSER solo en muy pocos enclaves de las provincias de Huesca (*De tú no me despido, pimpollo de albahaca fina*) y de Teruel (*Espérate allí hasta que venía el amo, a por tú*), mientras que tenían difusión más amplia en el ALEANR. En cambio, se ha podido constatar que otros aspectos dialectales mantienen incólume su área de uso, como es el caso del sistema referencial de los pronombres átonos de 3ª persona (León Zurdo 2017), el desplazamiento modal del imperfecto de subjuntivo (Pato 2003) o la pérdida de la *-d-* intervocálica (Estrada 2017).

Como hemos visto, no son pocos los rasgos morfosintácticos que han sido estudiados sobre datos del COSER. El resultado ha sido una mejor descripción y análisis de aspectos de variación que solo se describían de forma parcial o impresionista, tanto en lo relativo a sus límites geográficos como a sus factores lingüísticos condicionantes¹⁵. Esta investigación ha modificado notablemente la idea tradicional de que la variación gramatical en el español rural era casi inexistente, por lo que se ha obtenido un panorama mucho más complejo y también más ajustado a la realidad sobre la dialectología moderna del español, con no pocas repercusiones también para la historia de la lengua.

Así, la visión tradicional de las áreas dialectales del español europeo se ha visto enriquecida y complementada con aspectos relativos a la gramática. En algunas ocasiones, se corroboran áreas ya conocidas por la fonética –como el español meridional o el andaluz occidental– pero otras veces se postulan áreas nuevas que no se habían identificado claramente: el español oriental, el español occidental, el español noroccidental –y en contacto con el gallego y el asturiano–, el español septentrional, el español en contacto con el valenciano y de Murcia o el español en contacto con el vasco, entre otras (Fernández-Ordóñez 2016).

Finalmente, y es un aspecto que no suele señalarse, el COSER y los numerosos proyectos a él asociados han servido para formar a toda una serie de investigadores, tanto en España como fuera de ella, especialistas en dialectología, variación gramatical y sintaxis dialectal. El COSER ha servido de modelo, asimismo, en la creación de otros corpus orales del español, como el COLEM (*Corpus oral de la lengua española en Montreal*) y comienza poco a poco su extensión a tierras americanas (COSER-Cuba, COSER-Uruguay). Además, se emplea también como material en la enseñanza del español como lengua extranjera a alumnos de niveles avanzados (Sanmiguel Mariño 2011).

En el momento actual, solo un 13 % de sus materiales se ofrece en línea, y se espera seguir desarrollando el corpus hasta completar todas las provincias e islas, así como corregir los posibles errores en las transcripciones existentes. Por ello, el COSER tiene ante sí notables desafíos desde el punto de vista técnico para poder ser una herramienta aún más útil para la investigación gramatical. Retos pendientes son la mejora de la anotación morfosintáctica existente y su parseo o análisis sintáctico (*parsing*) –por constituyentes y dependencias–, así como la transcripción automática mediante

¹⁵ No podemos citar todos aquí. Otra dimensión en vías de estudio es la sintaxis de los posesivos, por el ejemplo, el posesivo tónico (Pato 2015), con valor habitualmente focal, o el adjetivo pospuesto a un adverbio, en lugar de un sintagma preposicional con núcleo pronominal (*detrás de ti > tuyo, tuya*) (Salgado & Bouzouita 2017).

reconocimiento del habla (*respeaking*) para, al menos, adelantar la ardua labor de transcripción de las encuestas que todavía no han sido abordadas.

REFERENCIAS BIBLIOGRÁFICAS

- ALCYL: Alvar, M. (1999): *Atlas lingüístico de Castilla y León*. Valladolid: Junta de Castilla y León, 3 vols.
- ALEA: Alvar, M. (1961-1973): *Atlas lingüístico y etnográfico de Andalucía*. Granada: Universidad de Granada/ CSIC, 6 vols.
- ALEANR: Alvar, M. (1979-83): *Atlas lingüístico y etnográfico de Aragón, Navarra y La Rioja*. Zaragoza: Institución Fernando el Católico/ CSIC, 12 vols.
- ALEcant: Alvar, M. (1995): *Atlas lingüístico y etnográfico de Cantabria*. Madrid: Fundación Marcelo Botín, 2. vols.
- ALECMan: García Mouton, P. & F. Moreno Fernández, *Atlas lingüístico (y etnográfico) de Castilla-La Mancha*. Alcalá de Henares: Universidad de Alcalá.
- ALEICan: Alvar, M. (1975): *Atlas lingüístico y etnográfico de las islas Canarias*. Madrid: Arco/ Libros, 3 vols.
- ALPI: Navarro Tomás, T. (dir.) (1962): *Atlas Lingüístico de la Península Ibérica*, vol. 1, *Fonética*. Madrid: CSIC. Véase también Pilar García Mouton (coord.), Inés Fernández-Ordóñez, David Heap, Maria Pilar Perea, João Saramago, Xulio Sousa, 2016, ALPI-CSIC [www.alpi.csic.es], edición digital de Navarro Tomás, Tomás (dir.), *Atlas Lingüístico de la Península Ibérica*, Madrid, CSIC.
- BARBIERS, S., L. CORNIPS & S. VAN DER KLEIJ (eds.). (2002): *Syntactic microvariation*. Amsterdam: Meertens-Instituut.
- BIBERAUER, T. (ed.) (2008): *The limits of syntactic variation*. Amsterdam/ Philadelphia: John Benjamins.
- BOUZOUTA, M. & E. PATO (2019): “*Antes había(n) pozos en el pueblo*. La pluralización del verbo *haber* existencial en español rural europeo”, *Revue de linguistique romane* 83, pp. 137-165.
- BUESA, T. & R. M. CASTAÑER (1994): “El pretérito perfecto simple en las hablas pirenaicas de Aragón y Navarra”, *Archivo de Filología Aragonesa* L, pp. 65-132.
- CAMUS BERGARECHE, B. & S. GÓMEZ SEIBANE (2015): “La diversidad del español en Álava: Sistemas pronominales a partir de las encuestas del COSER”, *Revista de Filología Española* XCV, pp. 279-206.
- CASTILLO LLUCH, M. & Á. OCTAVIO DE TOLEDO Y HUERTA (2016): “*Habemos muchos que hablamos español*: distribución e historia de la concordancia existencial en primera persona del plural”, in C. de Benito & Á. Octavio de Toledo y Huerta (eds.): *En torno*

a haber: *construcciones, usos y variación desde el latín hasta la actualidad*. Bern: Peter Lang, pp. 111-168.

CHAMBERS, J. K. (1995): *Sociolinguistic theory*. Oxford: Blackwell.

CHAMBERS, J. K. (2004): "Dynamic typology and vernacular universals", in B. Kortmann (ed.): *Dialectology meets typology*. Berlin/New York: Mouton de Gruyter, pp. 127-145.

CORBETT, G. C. (2006): *Agreement*. Cambridge: Cambridge University Press.

COLEM: Pato, E. (dir.) (2013-2020): *Corpus oral de la lengua española en Montreal*. Montreal: Université de Montréal <esp-montreal.jimdo.com>.

CORPES XXI: *Corpus del español del siglo XXI*. Madrid: RAE <<http://www.rae.es>>.

COSER: Fernández-Ordóñez I. (dir.) (2005-2019): *Corpus Oral y Sonoro del Español Rural*. Madrid: Universidad Autónoma de Madrid <corpusrural.es>.

CREA: *Banco de datos. Corpus de referencia del español actual*. Madrid: RAE <<http://www.rae.es>>.

DE BENITO MORENO, C. (2012): "The pronominal coding of the patient in reflexive indefinite agent constructions in Peninsular Spanish", *Journal of Portuguese Linguistics* 11(1), pp. 45-60.

DE BENITO MORENO, C. (2013): "(Esa tela) se la descose: la pronominalización del paciente en las impersonales reflejas del español peninsular", *Borealis* 2(2), pp. 129-157.

DE BENITO MORENO, C. (2015a): "Pero se escondíamos como las ratas: syncretism in the reflexive paradigm in Spanish and Catalan", *Isogloss* 1, pp. 95-127.

DE BENITO MORENO, C. (2015b): *Las construcciones con "se" desde una perspectiva variacionista y dialectal*. Tesis doctoral. Universidad Autónoma de Madrid.

DE BENITO MORENO, C. (2016): "La pronominalización en las construcciones existenciales con haber: ¿hay restricciones o no las hay?", in C. de Benito & Á. Octavio de Toledo y Huerta (eds.): *En torno a haber: construcciones, usos y variación desde el latín hasta la actualidad*. Bern: Peter Lang, pp. 209-237.

DE BENITO MORENO, C. & E. PATO (2015): "On the *de* + infinitive construction (*deísmo*) in Spanish", *Dialectologia* (Special issue V), pp. 29-51.

DE BENITO MORENO, C., J. PUEYO & I. FERNÁNDEZ-ORDÓÑEZ (2016): "Creating and designing a corpus of rural Spanish", *Proceedings of the 13th Conference on Natural Language Processing (KONVENS 2016)*. Bochum: Ruhr-Universität Bochum, pp. 78-83.

DEL BARRIO DE LA ROSA, F. (2018): *Espacio variacional y cambio lingüístico en español*. Madrid: Visor.

- ESTRADA ARRÁEZ, A. (2017): *The loss of intervocalic /d/ in the Iberian Peninsula*. Tesis doctoral. Universidad Autónoma de Madrid/Universidad de Friburgo.
- ESTRADA ARRÁEZ, A. (2018): “Factors involved in the evolution of intervocalic /d/ in the Iberian Peninsula: the case of frequency”, in F. Sánchez Miret & D. Recasens (eds.): *Production and Perception Mechanisms of Sound Change*. Múnich: Lincom Europa, pp. 145-156.
- FERNÁNDEZ-ORDÓÑEZ, I. (1993): “Leísmo, láismo y loísmo: estado de la cuestión”, in O. Fernández Soriano (ed.): *Los pronombres átonos*. Madrid: Taurus, pp. 63-96.
- FERNÁNDEZ-ORDÓÑEZ, I. (1994): “Isoglosas internas del castellano. El sistema referencial del pronombre átono de tercera persona”, *Revista de Filología Española* LXXIV, pp. 71-25.
- FERNÁNDEZ-ORDÓÑEZ, I. (1999): “Leísmo, láismo y loísmo”, in I. Bosque & V. Demonte (dirs.): *Gramática descriptiva de la lengua española*. Madrid: Espasa, vol. I, cap. 21, pp. 1317-1397.
- FERNÁNDEZ-ORDÓÑEZ, I. (2001): “Hacia una dialectología histórica. Reflexiones sobre la historia del leísmo, el láismo y el loísmo”, *Boletín de la Real Academia Española* LXXXI, pp. 389-464.
- FERNÁNDEZ-ORDÓÑEZ, I. (2006): “Del Cantábrico a Toledo: “El “neutro de materia” hispánico en un contexto románico y tipológico” (primera parte), *Revista de Historia de la Lengua Española* 1, pp. 67-118.
- FERNÁNDEZ-ORDÓÑEZ, I. (2007a): “Del Cantábrico a Toledo. El “neutro de materia” hispánico en un contexto románico y tipológico”, *Revista de Historia de la Lengua Española* 2, pp. 29-81.
- FERNÁNDEZ-ORDÓÑEZ, I. (2007b): “El neutro de materia en Asturias y Cantabria. Análisis gramatical y nuevos datos”, in I. Delgado Cobos & A. Puigvert Ocal (eds.): *Ex admiratione et amicitia. Homenaje a Ramón Santiago*. Madrid: Ediciones del Orto, pp. 395-434.
- FERNÁNDEZ-ORDÓÑEZ, I. (2007c): “Nuevas perspectivas en el estudio de la variación dialectal del español: El *Corpus Oral y Sonoro del Español Rural* (COSER)”, in D. Trotter (ed.): *Actes du XXIV Congrès de Linguistique et Philologie Romanes*. Tübingen: Niemeyer, vol. 3, pp. 29-44.
- FERNÁNDEZ-ORDÓÑEZ, I. (2009a): “Dialect grammar of Spanish from the perspective of the Audible Corpus of Spoken Rural Spanish (or Corpus Oral y Sonoro del Español Rural, COSER)”, *Dialectologia* 3, pp. 23-51.
- FERNÁNDEZ-ORDÓÑEZ, I. (2009b): “The development of mass/count distinctions in Indo-European languages”, in V. Bubenik, J. Hewson & S. Rose (eds.): *Grammatical change in Indo-European languages*. Amsterdam/Philadelphia: John Benjamins, pp. 55-68.

- FERNÁNDEZ-ORDÓÑEZ, I. (2010a): “La grammaire dialectale de l’espagnol à travers le Corpus oral et sonore de l’espagnol rural (COSER, *Corpus Oral y Sonoro del Español Rural*)”, *Corpus: La syntaxe de corpus/Corpus syntax* 9, pp. 81-114.
- FERNÁNDEZ-ORDÓÑEZ, I. (2010b): “New methods for the study of grammatical variation and the Audible Corpus of Spoken Rural Spanish”, in G. Aurrekoetxea & J. L. Ormaetxea (eds.): *Tools for Linguistic Variation*. Bilbao: Universidad del País Vasco, pp. 119-130.
- FERNÁNDEZ-ORDÓÑEZ, I. (2011): “Nuevos horizontes en el estudio de la variación gramatical del español: el *Corpus Oral y Sonoro del Español Rural*”, in G. Colón i Domènech & L. Gimeno Betí (eds.): *Noves tendències en la dialectologia contemporània*. Castellón de la Plana: Universitat Jaume I, pp. 173-203.
- FERNÁNDEZ-ORDÓÑEZ, I. (2012): “Dialect areas and linguistic change: Pronominal paradigms in Ibero-Romance dialects from a cross-linguistic and social typology perspective”, in G. de Vogelaer & G. Seiler (eds.): *The Dialect Laboratory. Dialects as a testing ground for theories of language change*. Amsterdam/Philadelphia: John Benjamins, pp. 73-106.
- FERNÁNDEZ-ORDÓÑEZ, I. (2015): “*Mucha trabajo*: sincretismo femenino en los cuantificadores evaluativos de Cantabria”, in *Studium Grammaticae. Homenaje al profesor José Antonio Martínez*. Oviedo: EdiUNo, pp. 337-349.
- FERNÁNDEZ-ORDÓÑEZ, I. (2016): “Dialectos del español peninsular”, in J. Gutiérrez Rexach (ed.): *Enciclopedia lingüística hispánica*. London/New York: Routledge, vol. 2, pp. 387-404.
- FERNÁNDEZ-ORDÓÑEZ, I. (2019): “Mass/count distinctions in Ibero-Romance dialects”, in Á. Gallego (ed.): *The syntactic variation of Spanish dialects*. Oxford: Oxford University Press, pp. 60-108.
- GÓMEZ SEIBANE, S. (2017): “Español en contacto con la lengua vasca: datos sobre la duplicación de objetos directos posverbiales”, in A. Palacios (ed.): *Variación y cambio lingüístico en situaciones de contacto*. Madrid/Frankfurt: Iberoamericana/Vervuert, pp. 143-159.
- HEAP, D. (2006): “Secuencias «invertidas» de clíticos: un cambio (¿?) en tiempo real”, in J. J. de Bustos Tovar & J. L. Girón Alconchel (eds.): *Actas del VI Congreso Internacional de Historia de la Lengua Española*. Madrid: Arco/Libros, vol. I, pp. 785-798.
- HEAP, D. & E. PATO (2012): “Plurales anómalos en los dialectos y en la historia del español”, in E. Montero Cartelle & C. Manzano Rovira (eds.): *Actas del VIII Congreso Internacional de Historia de la Lengua Española*. Santiago de Compostela: AHLE/ Meubook, vol. 1, pp. 829-840.
- KLEIN-ANDREU, F. (2000): *Variación actual y evolución histórica: los clíticos le/s, la/s, lo/s*. Múnich: Lincom Europa.

- KORTMANN, B. (1999): "Typology and dialectology", in B. Caron (ed.): *Proceedings of the 16th International Congress of Linguists*. Amsterdam: Elsevier Science, Multimedia CD.
- KORTMANN, B. (2004a): "Why dialect grammar matters", *The European English Messenger* XIII, pp. 24-29.
- KORTMANN, B. (ed.) (2004b): *Dialectology meets typology. Dialect grammar from a cross-linguistic perspective*. Berlin/New York: Mouton de Gruyter.
- LARA BERMEJO, V. (2018): "El superlativo absoluto en el español peninsular del siglo XX", *RILCE* 34 (1), pp. 225-239.
- LAPESA, R. (1968): "Sobre los orígenes y evolución del leísmo, laísmo y loísmo", in K. Baldinger (ed.): *Festschrift Walther von Wartburg*. Tübingen: Max Niemeyer, pp. 523-551.
- LEÓN ZURDO, O. (2017): "Case variation in unstressed third person pronouns in the Linguistic Atlas of the Iberian Peninsula", *Dialectologia* 18, pp. 43-72.
- LLORENTE MALDONADO DE GUEVARA, A. (1965): "Algunas características lingüísticas de La Rioja en el marco de las hablas del valle del Ebro y de las comarcas vecinas de Castilla y Vasconia", *Revista de Filología Española* 48(3/4), pp. 321-350.
- MONDÉJAR, J. (1970): *El verbo andaluz. Formas y estructuras*. Madrid: CSIC.
- OCTAVIO DE TOLEDO Y HUERTA, Á. (2016): "Sin CORDE pero con red: *algotras* fuentes de datos", *Revista Internacional de Lingüística Iberoamericana* 28, pp. 19-47.
- PATO, E. (2003): *La sustitución del imperfecto de subjuntivo por el condicional simple y el imperfecto de indicativo en el castellano septentrional peninsular. (Estudio de variación dialectal)*. Tesis doctoral. Universidad Autónoma de Madrid.
- PATO, E. (2004a): "Los perfectos fuertes analógicos en español", in M. Trinidad (ed.): *Actas del Congreso Internacional "APLEx 2004". Patrimonio Lingüístico Extremeño* (Cáceres, 4-6 noviembre de 2004). Cáceres: Editora Regional (CD Rom).
- PATO, E. (2004b): *La sustitución de cantara/cantase por cantarí/cantaba (en el castellano septentrional peninsular)*. Madrid: Universidad Autónoma de Madrid.
- PATO, E. (2010a): "Los adverbios *ahora* y *ahora*: Dos orígenes, un mismo resultado", *Revista de Historia de la Lengua Española* 5, pp. 167-173.
- PATO, E. (2010b): "Linguistic levelling in Spanish: The analogical strong preterites", *Canadian Journal of Linguistics* 55(2), pp. 209-225.
- PATO, E. (2012a): "*Cantábamos* por *cantábamos*: forma 'etimológica' del español rural", *Archivo de Filología Aragonesa* 68, pp. 219-236.
- PATO, E. (2012b): "Nivelación lingüística y simplificación: El uso de preposición + *tú* en la historia de la lengua", in E. Montero Cartelle & C. Manzano Rovira (eds.): *Actas*

del VIII Congreso Internacional de Historia de la Lengua Española. Santiago de Compostela: AHLE/Meubook, vol. 1, pp. 1029-1039.

- PATO, E. (2012c): “Variación dialectal y análisis estadístico: Formas indicativas por subjuntivas en español rural”, in E. Pato & J. Rodríguez Molina (eds.): *Estudios de filología y lingüística españolas. Nuevas voces en la disciplina*. Bern: Peter Lang, pp. 93-133.
- PATO, E. (2013): “Sobre la forma *muncho*”, *Estudios de Lingüística. Universidad de Alicante* 27, pp. 329-342.
- PATO, E. (2015): “El posesivo (antepuesto) tónico en español: ¿fenómeno de foco y contraste?”, *Dialectología* 14, pp. 47-73.
- PATO, E. (2016): “La pluralización de *haber* en español peninsular”, in C. de Benito & Á. Octavio de Toledo y Huerta (eds.): *En torno a haber: construcciones, usos y variación desde el latín hasta la actualidad*. Bern: Peter Lang, pp. 357-391.
- PATO, E. (2018a): “Semejante adjetivo es *semejante*. Sus valores en español actual”, *Estudios filológicos* 61, pp. 59-74.
- PATO, E. (2018b). “*Queriba una cosa y traíba otra*. Los pretéritos imperfectos ‘analógicos’ en español”, *Philologica Jassyensia* XIV(2), pp. 83-100.
- PATO, E. & P. O’NEILL (2013): “Los gerundios ‘analógicos’ en la historia del español (e iberorromance)”, *Nueva Revista de Filología Hispánica* LXI(1), pp. 1-27.
- PATO, E. & X. VIEJO FERNÁNDEZ (2017): “*Mu(n)cho guapa y mu(n)cho bien: mu(n)cho* como cuantificador ‘pleno’ de adjetivos y adverbios en castellano y en asturiano”, *Bulletin of Spanish Studies* 94(1), pp. 1-23.
- SALGADO, H. & M. BOUZOUITA (2017): “El uso de las construcciones de adverbio locativo con pronombre posesivo en el español peninsular: un primer acercamiento diatópico”, *Zeitschrift für romanische Philologie* 133(3), pp. 766-794.
- SANMIGUEL MARIÑO, A. (2011): “Reseña: *Corpus Oral y Sonoro del Español Rural* (COSER)”, *Boletín de ASELE* 44, pp. 58.