

# VI CONGRESO INTERNACIONAL DE CORPUS DIACRÓNICOS EN LENGUAS IBERORROMÁNICAS

VENEZIA, 5-7 DE OCTUBRE DE 2022

## LIBRO DE RESÚMENES



Università  
Ca'Foscari  
Venezia

Dipartimento di Studi Linguistici  
e Culturali Comparati

Con la colaboración de:



*Asociación de Historia de la Lengua Española*

**AISPI** ASSOCIAZIONE  
ISPANISTI  
ITALIANI



M i l á n

/10.13039/501100011033/ FEDER  
(Una manera de hacer Europa)

Secretaría administrativa:



Fondazione  
Università  
Ca'Foscari

## ÍNDICE POR AUTORES

Bejarano Bejarano, Daniel Eduardo.....	61
Benito Moreno, Carlota de .....	6
Blas Arroyo, José Luis .....	29
Bonilla, Johnatan E.....	53
Caballero Gómez, Víctor; Ricardo Pichel y Diego Sánchez Sierra .....	16
Calderón Campos, Miguel y Gael Vaamonde .....	51
Castillo Lluch, Mónica y Andrea Escobar Castillo .....	19
Cornillie, Bert y Giulia Mazzola .....	58
Corral Esteban, Avelino .....	42
Costa León, Piero; Dunia Hourani, Matteo De Beni y Elisa Sartor .....	33
Enrique-Arias, Andrés y Katharina Gerhalter .....	57
Enrique-Arias, Andrés; Marina Gomila Albal y Florencio del Barrio.....	66
Enrique-Arias, Andrés; Silvia Perpiñán e Ileana Paul.....	22
Espejel, Marina .....	25
Fernández de Molina Ortés, Elena.....	26
Ferrari, Laura D. ....	52
Fiéis, Alexandra .....	2
Flores, Cristina y Esther Rinke.....	21
Gago-Jover, Francisco.....	65
Garachana, Mar y Álvaro Octavio de Toledo y Huerta .....	13
García Aguiar, Livia C.....	15
Heredia Mantis, María.....	41
Jiménez, María Isabel.....	23
Julià Luna, Carolina; Carmen Quijada Van den Berghe y Assumpció Rost Bagudanch .....	60
Junquera Martínez, Alejandro .....	30
Kabatek, Johannes y Manuela Crivelli .....	35
López Mora, Pilar.....	14
López Serena, Araceli y Santiago del Rey Quesada.....	49
Mabille, Charles .....	34
Marcet Rodríguez, Vicente J. y M. <sup>a</sup> Nieves Sánchez González de Herrero.....	63
Martins, Ana Maria.....	47
McCarley, Gemma .....	32
Medina López, Javier.....	39
Miguel Franco, Ruth y David Sánchez .....	41

Molenaers, Marie .....	20
Moral del Hoyo, M. <sup>a</sup> Carmen .....	7
Nevot Navarro, Manuel .....	64
Peláez Santamaría, Salvador .....	38
Pérez Saldanya, Manuel.....	5
Pichel, Ricardo y Xavier Varela Barreiro .....	18
Pinto, Clara.....	12
Pons Bordería, Salvador.....	52
Primerano, Antoine .....	56
Ramos Remedios, Emiliana.....	8
Ribeiro, Maria .....	11
Robles Ávila, Sara .....	54
Rodríguez Guerra, Alexandre .....	10
Rodríguez Molina, Javier .....	35
Rosemeyer, Malte.....	3
Salameh Jiménez, Shima .....	28
Santos Costa, Ana; Ana Paula Soares y Montserrat Comesaña.....	40
Sanz Martin, Blanca Elena.....	36
Sentí, Andreu y Miriam Bouzouita.....	46
Sousa Fernández, Xulio.....	62
Torrens Álvarez, María Jesús y Hiroto Ueda.....	9
Torres Latorre, Aina.....	45
Torres Soler, Julio.....	24

## SESIONES PLENARIAS

---

### Conferencia inaugural Miércoles 5, 13:30

---

**ALEXANDRA FIÉIS**  
*Universidade Nova de Lisboa*

#### **Corpus linguistics: a methodological basis for diachronic syntax studies**

In this presentation, we aim at describing some aspects of Portuguese morphosyntax from a diachronic perspective, based on the description of written corpora, with the purpose of better understanding grammatical phenomena that are subject to variation (e.g. the syntax of transitive and intransitive constructions, cliticization and the functional domain, adverbial subordination, multifunctional words and functional structure, and nonfinite complementation), while trying to understand the advantages of this type of methodology.

By doing so and by understanding the changes that occurred both at the morphosyntactic and the syntactic-semantic levels in different stages of Portuguese, we aim also to contribute to the understanding of the mechanisms of language variation and change.

---

## Conferencia plenaria Jueves 6, 18:30

---

**MALTE ROSEMEYER**  
*Freie Universität Berlin*

### **Apuntes sobre la utilidad de las variables formadas a base de datos agregados para la descripción de los cambios históricos**

La presente charla adopta una perspectiva metodológica sobre la descripción de los cambios históricos y, en particular, la cuestión de cómo se puede modelar cambios lingüísticos que surgen en la interacción utilizando datos históricos cuantitativos. Muchas teorías del cambio semántico asumen que dichos cambios no son aleatorios, sino que surgen de estrategias comunicativas específicas en interacción, como la expresividad (Detges, Waltereit 2002), la organización del discurso (Rosemeyer, Grossman 2017; 2021) o la economía (p. ej., Haspelmath 1999). Sin embargo, cuando se aplican a los datos históricos reales, estas hipótesis sobre el cambio semántico son difíciles de verificar. Esto se debe al hecho de que los documentos escritos que han sobrevivido hasta hoy casi nunca representan datos de una interacción lingüística espontánea real. Sostengo que para poder evaluar la relevancia de cada una de estas hipótesis para los cambios históricos, necesitamos desarrollar métodos que nos permitan medir hasta qué punto los datos históricos analizados reflejan la interacción espontánea en el período en cuestión.

Sobre la base de estudios previos (Szmrecsanyi 2016; Rosemeyer 2019; Rosemeyer, Cornillie 2021; Rosemeyer, Becker, en prensa), propongo que al aplicar el análisis de registro (Biber, Finegan 1997; Biber, Finegan 2004) a los corpus diacrónicos, es posible describir el grado al que estos textos se aproximan a la interacción hablada real en el momento de la composición de la obra. Por consiguiente, este análisis demuestra la utilidad de las variables formadas a base de datos agregados para la lingüística de corpus diacrónica. En particular, se describe el grado de formalidad de cada texto en un corpus a base de las frecuencias agregadas de fenómenos asociados al discurso informal, como, por ejemplo, verbos que tienen una semántica intersubjetiva, la deixis local y temporal, y los marcadores de discurso. De este modo, las frecuencias de uso agregadas de estos fenómenos se convierten en una escala monotónica que modela el fenómeno abstracto que es el grado de formalidad de un texto.

Centrándome en el desarrollo de construcciones interrogativas y el presente perfecto en el portugués de Brasil, y las perífrasis verbales en el español europeo, demuestro la relevancia de este enfoque. En particular, la incorporación de medidas globales de formalidad relacionadas con el registro en los análisis diacrónicos cuantitativos del cambio lingüístico permite determinar qué cambios construccionales pueden describirse como originados en la interacción informal y cuáles parecen resultar de “cambios desde arriba” marcados estilísticamente (Meyerhoff 2006: 222-225). Además, este enfoque se comparará a intentos recientes de establecer catálogos de las tradiciones discursivas representadas en un corpus a base de un análisis inductivo de datos agregados (Mazzola, De Pascale, Rosemeyer, en prensa).

### **Referencias bibliográficas**

Biber, D.; Finegan, E. (1997). “Diachronic Relations among Speech-Based and Written Registers in English”. Nevalainen, T.; Kahlas-Tarkka, L. (eds.), *To*

*Explain the Present: Studies in the Changing English Language in Honor of Matti Rissanen.* Helsinki: Société Néophilologique, 253-275.

- Biber, D.; Finegan, E. (2004). "Historical drift in three English genres". Sampson, G; McCarthy, D. (eds.), *Corpus Linguistics: Readings in a Widening Discipline*. London: Continuum, 67-77.
- Detges, U.; Waltereit, R. (2002). "Grammaticalization vs. reanalysis: a semantic-pragmatic account of functional change in grammar". *Zeitschrift für Sprachwissenschaft* 21(2), 151-195.
- Haspelmath, M. (1999). "Explaining Article-Possessor Complementarity: Economic Motivation in Noun Phrase Syntax". *Language*, 75(2), 227-243.
- Mazzola, G.; De Pascale, S.; Rosemeyer, M. (en prensa). "Nuevas herramientas en la lingüística diacrónica: tradiciones discursivas y lingüística computacional". Cornillie, B.; Mazzola, G.; Thegel M. (eds.), *La Tradicionalidad Discursiva y la Lingüística de Corpus: conceptos y aplicaciones*. Madrid, Frankfurt: Iberoamericana Vervuert.
- Meyerhoff, M. (2006). *Introducing Sociolinguistics*. London, New York: Routledge.
- Rosemeyer, M. (2019). "Actual and apparent change in Brazilian Portuguese *wh*-interrogatives". *Language Variation and Change*, 31(2), 165-191.
- Rosemeyer, M.; Becker M. (en prensa). "The Brazilian Portuguese present perfect: from nominal to verbal plurality". Kellert, O.; Rosemeyer M. (eds.), *Indefinites in Romance and Beyond*. Berlin: Language Science Press.
- Rosemeyer, M.; Cornillie B. (2021). "Syntactic elaboration: The case of auxiliation in the communicative distance". Charla en el *54 Annual Meeting of the Societas Linguistica Europae* (SLE, 3 de septiembre de 2021).
- Rosemeyer, M.; Grossman E. (2017). "The road to auxiliariness revisited: the grammaticalization of FINISH anteriors in Spanish". *Diachronica*, 34(4), 516-558.
- Rosemeyer, M.; Grossman E. (2021). "Why don't grammaticalization pathways always recur?". *Corpus Linguistics and Linguistic Theory*, 17(3), 653-681.
- Szmrecsanyi, B. (2016). "About text frequencies in historical linguistics: Disentangling environmental and grammatical change". *Corpus Linguistics and Linguistic Theory*, 12(1), 153-171.

---

## Conferencia plenaria Viernes 7, 12:30

---

**MANUEL PÉREZ SALDANYA**  
*Universitat de València*

### **Al voltant de 'aquí' i dels corpus textuais informatitzats del català**

En aquesta presentació ens centrem en els principals corpus textuais informatitzats del català, tant els de caràcter diacrònic (CICA, CIGCMod, CIVAl, DTCA) com sincrònic (CTILC). Delimitarem les seves característiques i tractarem de donar compte de les seves forteses i, si en tenen, de les seves limitacions. Per a analitzar i exemplificar aquests aspectes, ens centrarem en l'estudi d'una classe de paraules transversal del català; concretament, en la classe dels demostratius, integrada per determinants (p. ex. *est(e)/aquest*), pronoms (p. ex. *açò*), elements pròxims a l'article (*ço*) i adverbis (p. ex. *aquí, hi*).

Fent servir els corpus textuais del català, tractarem de delimitar si és possible datar i explicar els processos de canvi lingüístic a què s'han vist sotmesos i que afecten aspectes morfològics, sintàctics, pragmàtics i sociolingüístics; per exemple: a) la formació d'un sistema de dixi ternari a partir del sistema binari del llatí tardà; b) l'adopció posterior d'un sistema binari en la major part del català; c) la bifurcació dialectal que es produeix entre els parlars amb formes de determinants reforçades (*aquest*) i formes simples (*est(e)* i *eix(e)*); d) la decadència del demostratiu *ço* i la substitució per l'article *lo* en alguns dels seus usos; i e) la repercussió de la codificació fabriana en les formes documentades en els textos, especialment a partir de la publicació de la *Gramàtica catalana* (1918/1933).

### **Corpus textuais informatitzats**

CICA = Torruella J. (dir.); Pérez Saldanya M.; Martines J. *Corpus Informatitzat del Català Antic*. <http://cica.cat/>.

CIGCMod = Martines, J.; Martines V. (eds.). *Corpus Informatitzat de la Gramàtica del Català Modern* (en procés). ISIOC-IVITRA, Universitat d'Alacant.

CITLC = Rafel, J. (dir.). *Corpus Informatitzat Textual de la Llengua Catalana*. Institut d'Estudis Catalans. <https://ctilc.iec.cat/>.

CIVAL = Acadèmia Valenciana de la Llengua. *Corpus Informatitzat del Valencià*. <http://cival.avl.gva.es/>.

DTCA = Universitat de Barcelona. *Diccionari de Textos Catalans Antics*. <http://www.ub.edu/diccionari-dtca/>.

---

**Conferencia de clausura**  
**Viernes 7, 18:30**

---

**CARLOTA DE BENITO MORENO**  
*Universität Zurich*

***Ojalá tener todos los datos. Cuestiones de método en el uso de las redes sociales digitales para el estudio de la microdiacronía***

Es ya casi un tópico subrayar el interés del lenguaje juvenil y el coloquial para el estudio del cambio lingüístico, por la velocidad con que se producen determinadas renovaciones (especialmente léxicas, aunque no exclusivamente) en ámbitos como la intensificación o la negación en estas variedades. Gracias a la generalización de internet y el advenimiento de las redes sociales digitales —ya no tan recientes—, se ha abierto una nueva posibilidad de estudiar dichos cambios, dado que las condiciones comunicativas de las redes sociales digitales favorecen la escritura de formas de expresión propias de la inmediatez comunicativa y, por tanto, facilitan enormemente su estudio. Sin embargo, el uso de los textos digitales como corpus lingüístico no está exento de dificultades metodológicas, algunas ya viejas conocidas de la filología y la lingüística de corpus y otras, novedosas. Serán precisamente las cuestiones de método las que nos guíen en la exposición de este estudio sobre la microdiacronía de un cambio lingüístico que, por ahora, parece afectar únicamente al español coloquial, a saber, la pérdida de restricciones sintácticas experimentada por la forma *ojalá*. Esta ¿interjección? ha pasado de únicamente introducir oraciones completivas en subjuntivo a permitir también oraciones no finitas con infinitivo (*Ojalá estar viajando constantemente*, tuit, 26/04/2017), gerundio (*Ojalá mi abuela haciéndome una coleta de las de no poder cerrar los ojos para ir a misa ahora*, tuit, 27/01/2013), construcciones absolutas con participio (*Ojalá un gobierno permanentemente en funciones liderado por Chenoa*, tuit, 25/06/2016) o, incluso, construcciones sin verbo, típicamente con interpretación existencial (*Lo digo desde ya: ojalá unas terceras elecciones*, tuit, 25/06/2016). Hablaremos del origen del cambio y de su difusión, prestando atención también a los hablantes que las producen y las conexiones entre ellos.



## COMUNICACIONES

Miércoles, 5 de octubre de 2022

Aula Baratto

Sesión Diseño de Corpus (1)

15:00 – 16:30

### El cierre de las vocales velares finales en el castellano norteño medieval

M.<sup>a</sup> Carmen Moral del Hoyo  
*Universidad de Cantabria*

La visión tradicional sobre la fluctuación del vocalismo velar final *-o/-u* en castellano muestra una alternancia poco testimoniada ya en la etapa de orígenes del español (siglos X y XI) (Menéndez Pidal 1968, López García 1977). La excepción la constituirían algunos testimonios esporádicos en la denominada Cantabria primitiva, en la que persistirían por el arcaísmo lingüístico achacado a esta región (y en la que, sin embargo, perviven actualmente).

Así, los estudios que han intentado trazar el devenir histórico de este vocalismo velar cerrado [-u] se han dirigido a explicar la situación actual de su distinto tratamiento únicamente en el área asturleonese, de la que se considera prototípico en las descripciones dialectales presentes y pasadas. En ellos se dirime si la resistencia de *-u* es mero arcaísmo latino, si es posible que adquiriese nuevos valores morfológicos en romance que favorecieran su conservación y la relación de todo ello con la expresión de la (dis)continuidad a través del denominado “neutro de materia” (Neira 1991, Fernández-Ordóñez 2006, 2007, Viejo 2018). Pero en los estudios históricos sobre estos ámbitos dialectales, ha sido común integrar indiscriminadamente en la descripción casuística las voces que atestiguan este cierre de la vocal final independientemente de las condiciones gramaticales y, por tanto, sin reparar en las posibles implicaciones morfosintácticas.

Así pues, no se había acometido un estudio sistemático de la presencia de *-u* en documentación castellana medieval, trabajo que se plantea en Moral del Hoyo (2020). La disponibilidad de valiosos corpus como CORHEN y CODEA+, en los que los documentos están datados cronológica y, sobre todo, tópicamente de forma segura y fidedigna, permite acometer esta revisión y sistematización de *-u* en castellano medieval, desde una perspectiva intradialectal y, con ello, establecer su papel en la dinámica histórica de la marcación de la (dis)continuidad que podemos ahora ir enriqueciendo con nuevos datos.

### Referencias bibliográficas

- CODEA+: Grupo de Investigación Textos para la Historia del Español (GITHE) (2015). *Corpus de documentos españoles anteriores a 1800*. [corpuscodea.es](http://corpuscodea.es).
- CORHEN = Torrens Álvarez, M.<sup>a</sup> J. (dir. y ed.). *Corpus Histórico del Español Norteño*. <http://corhen.es/>.
- Fernandez-Ordóñez, I. (2006). “Del Cantábrico a Toledo: el *neutro de materia* hispánico en un contexto románico y tipológico”. *Revista de Historia de la Lengua Española*, 1, 67-118.

- Fernandez-Ordóñez, I. (2007). "Del Cantábrico a Toledo: el *neutro de materia* hispánico en un contexto románico y tipológico (continuación)". *Revista de Historia de la Lengua Española*, 2, 29-81.
- López García, Á. (1977). "Simetría, asimetría y disimetría en el vocalismo iberorrománico". *Archivum*, 27-28, 293-334
- Menéndez Pidal, R. (1968). *Orígenes del español*. Madrid: Espasa-Calpe.
- Moral del Hoyo, M.<sup>a</sup> C. (2020). "(Dis)continuidad y vocal final –u en castellano: de Orígenes (y de orígenes) al siglo XIII". Fernández-Ordóñez, I. (coord.), *El legado de Ramón Menéndez Pidal (1869-1968) a principios del siglo XXI*. CSIC, Anejos de la RFE, vol. 1, 293-329.
- Viejo Fernández, X. (2018). "La morfologización de las armonizaciones vocálicas en el centro de Asturias: innovación y escisión protorromance". *Verba*, 19, 193-224.

---

### De nuevo sobre los inicios de la escritura castellana a raíz de algunos documentos hebreos del Corpus del Español Norteño

Emiliana Ramos Remedios  
*Universidad del País Vasco*

Para explicar el proceso de gestación de la escritura romance se ha planteado la posibilidad de que algunas minorías sociales como judíos, francos o ciertas órdenes religiosas hayan sido el motor de este desarrollo (Hernández 2009). No obstante, la propia heterogeneidad de los fondos en los que estos primeros textos romances se encuentran –tanto entornos rurales como urbanos– y un estudio pormenorizado y comparativo de documentación procedente de diversas áreas como la que aporta el *Corpus del Español Norteño* (CORHEN) inclinan a pensar en que es el ámbito urbano el que motiva que la escritura en lenguas vernáculas y, por tanto, la castellana eclosionen, como defiende Torrens Álvarez (2013).

En esta línea y volviendo sobre algunos documentos como los hebreos de Aguilar de Campoo (fechados en 1187, 1219 y 1220) u otros documentos castellanos particulares –burgaleses, cántabros, palentinos y vallisoletanos– incluidos en el corpus *CORHEN* y relacionados con judíos, reflexionaremos de nuevo acerca de si estos grupos sociales pudieron tener algún peso en la consolidación del proceso de escritura en romance, y más especialmente en romance castellano, sin dejar de lado, por otra parte, la posibilidad de que la escritura aljamiada, que puede estar tras algunos de esos textos, seleccionara una norma lingüística diferente a la de los textos romances castellanos (Del Barco 2011; Alfonso y del Barco 2021).

#### Referencias bibliográficas

- Alfonso, E.; del Barco, J. (2021). *Translating the Hebrew Bible in Medieval Iberia*. Oxford, Bodleian Library, MS Hunt. 268. Leiden/Boston: Brill.
- Del Barco, Javier (2011). "Estudio lingüístico de glosas romances en un manuscrito exegético hebreo del siglo XIII". *Romance Philology*, 65(2), 173-193.
- CORHEN = Torrens Álvarez, M.<sup>a</sup> J. *Corpus Histórico del Español Norteño*. <http://www.corhen.es>.

- Hernández, F. J. (2009). "The Jews and the Origins of Romance Script in Castile: A New Paradigm". *Medieval Encounters*, 15, 259-306.
- Torrens Álvarez, M.<sup>a</sup> J. (2013). "Sobre la apuesta por la escritura romance en los monasterios burgaleses: Oña vs. Las Huelgas". *Cuadernos del Instituto de Historia de la Lengua*, 8, 321-342.

---

## **CORHEN en LYNEAL: herramientas para el estudio microdialectal del castellano medieval norteño**

María Jesús Torrens Álvarez / Hiroto Ueda  
CSIC / Universidad de Tokio

El *Corpus Histórico del Español Norteño* (CORHEN), dirigido por María Jesús Torrens, está formado por la documentación original de carácter particular escrita en los territorios de las actuales provincias de Burgos y Palencia durante los siglos X al XIII. Editada en una doble presentación, paleográfica y crítica o normalizada, según los criterios de la Red CHARTA, esta documentación permite estudiar el paso del latín al romance escrito, así como llevar a cabo estudios de microdialectología histórica y poner de manifiesto la complejidad dialectal de un área limitada en un periodo igualmente acotado, con resultados que contradicen la supuesta homogeneidad que se ha atribuido a la variedad castellana desde sus orígenes.

Los análisis de los datos lingüísticos son posibles gracias al sistema LYNEAL (Letras y Números en Análisis Lingüísticos), diseñado por Hiroto Ueda, que permite buscar formas y patrones mediante el empleo de expresiones regulares; realizar el filtrado y la combinación de atributos; efectuar diversos tipos de análisis cuantitativos o estadísticos (frecuencia absoluta, relativa, normalizada, probabilística, etc.), así como multivariantes; y representar dichos resultados en distintos formatos de gráficos, tablas y mapas.

Así, por ejemplo, y es este el aspecto que más nos interesará presentar, podrá visualizarse en mapas sucesivos la combinación de la distribución geográfica y diacrónica, en cortes cronológicos por periodos de años establecidos libremente por el usuario del programa, mapas que pondrán de manifiesto los cambios diatópicos habidos a lo largo del tiempo.

El objetivo de la comunicación será ofrecer distintas muestras de las posibilidades de estas herramientas de explotación lingüística aplicadas al corpus CORHEN, con especial atención, como se acaba de mencionar, a la visualización en mapas diacrónicos. Para ello, se analizarán diversos fenómenos de variación gráfica, como <ce, i> / <çe, i>; fonética, como la asimilación *enna* / *en la*; morfológica, como los demostrativos *este* / *aqueste*; o léxica, como el reparto de los anafóricos *sobredicho* / *susodicho* / *avandicho*, o de palabras como *flumen* / *rivus*.

La comparación de los resultados de CORHEN con los proporcionados por otros corpus más amplios en su cronología y geografía, pero con una representación por regiones y periodos necesariamente más limitada, permitirá mostrar el interés de la microdialectología histórica, que devuelve una imagen mucho más compleja de la realidad.

Aula Berengo  
**Sesión Sintaxis Histórica (1)**  
15:00 – 16:30

**A posición do verbo finito no galego medieval: V1, V2... Vn**

Alexandre Rodríguez Guerra  
*Universidade de Vigo*

Diferentes estudos analizaron as linguas románicas na súa diacronía para comprobar se evidenciaban un comportamento de V2 na orde dos seus constituíntes oracionais. Entre outros están Ribeiro (1995), Martins (2014) ou Medeiros (2015) para o portugués medieval, a monografía de López Izquierdo/Castillo Lluch (2015) para – esencialmente- o castelán medieval e, dende unha perspectiva catalá e panrománica, Batllori (2015) ou Sitaridou (2012). Pero aínda non contamos cun estudo específico sobre o caso galego (cfr. *BILEGA*), malia que si existe, por exemplo, a publicación de Ledgeway (2017) sobre un texto latinovulgar galaico tan importante como o *Itinerarium Egeriae*.

A partir do noso corpus oracional do galego medieval (séculos XIII-XV), nunha versión máis actualizada cá descrita en Rodríguez Guerra (2008-2009), comprobamos o peso das distintas posicións V2, V1 e  $V \geq 3$ , a (a)simetría entre os diferentes contextos oracionais, así como a repercusión doutros parámetros como son a tipoloxía textual e a cronoloxía. Alén diso, tamén analizamos a posible influencia tanto da natureza sintáctica dos constituíntes coma a tipoloxía verbal. Estudamos, igualmente, a repercusión dos esquemas sintácticos, as grandes tipoloxías relacionais e o carácter (non) argumental dos constituíntes léxicos (sen desbotarmos completamente información sobre os clíticos antepostos ó verbo).

Os resultados desta investigación permiten verificar as liñas mestras que orientan as escollas de V1, V2 e  $V \geq 3$  nos textos galegos medievais. Grazas ás características do corpus manexado, esas liñas son suficientemente reveladoras e posibilitarán unha rigorosa análise contrastiva co resto de linguas (iberor)romances, o que facilitará o agromo das especificidades e, dende unha perspectiva máis xeral, un deseño moito máis completo do mapa románico neste punto da súa sintaxe.

### Referencias bibliográficas

Batllori, M. (2015). "Análisis comparativo de las construcciones de dislocación a la derecha y su incidencia en el orden de palabras del español, catalán y portugués medievales". López Izquierdo, M.; Castillo Lluch, M. (coords.), *El orden de palabras en la historia del español y otras lenguas iberorromances*. Madrid: Visor Libros, 83-110.

*BILEGA* = García Gondar, F.; Mariño Paz, R. (dirs.) (1994-). *Bibliografía Informatizada da Lingua Galega*. Santiago de Compostela: CRPIH.  
<https://bernal.cirp.gal/ords/f?p=107:1>.

Ledgeway, A. (2017). "Late Latin Verb Second: The Sentential Word Order of the *Itinerarium Egeriae*". *Catalan Journal of Linguistics*, 16, 163-216.

López Izquierdo, M.; Castillo Lluch, M. (coords.) (2015). *El orden de palabras en la historia del español y otras lenguas iberorromances*. Madrid: Visor Libros.

- Martins, A. M.<sup>a</sup> (2014). "An Argument Against Verb-Second in Old Portuguese". Fiéis, A.; Lobo, M.<sup>a</sup>; Madeira, A. (eds.), *O Universal e o Particular: Uma vida a comparar, Homenagem a Maria Francisca Xavier*. Lisboa: Colibri, 207-216.
- Medeiros, C. Salgado Lacerda (2015). "Ordem V2 no português arcaico: revisitando a problemática da periodização". *Labor Histórico*, 1/2, 108-123.
- Ribeiro, I. (1995). *A sintaxe da ordem no português arcaico: o efeito V2*. Tese de Doutoramento. Universidade Estadual de Campinas.
- Rodríguez Guerra, A. (2008-2009). "A oración en galego medieval (AOGM): O banco de datos sintácticos medievais do ILG". *Cadernos de Lingua*, 30/31, 123-150.
- Sitaridou, I. (2012). "A comparative study of word order in Old Romance". *Folia Linguistica*, 46/2, 553-604.

---

## Aspectos do percurso evolutivo dos verbos *ser* e *estar* em português medieval

Maria Ribeiro  
*Universidade Nova de Lisboa*

Em português contemporâneo, à semelhança do que sucede em outros sistemas ibero-românicos, são associados aos verbos *ser* e *estar* valores diferenciados que se caracterizam por marcar a oposição entre uma relação perspectivada como 'inerente' ou 'permanente' e uma relação 'temporária' ou 'transitória'. Porém, esta distribuição de valores nem sempre esteve consolidada na história da língua portuguesa, observando-se, em português medieval, uma sobreposição parcial dos valores semânticos associados a estes verbos. Em específico, em fases anteriores da língua portuguesa, constata-se que o verbo *ser*, paradigma supletivo resultante da fusão dos verbos latinos *esse* 'ser' e *sedere* 'estar sentado', era utilizado em contextos em que atualmente apenas se admitiria o verbo *estar* (< *stare* 'estar sentado').

Partindo da análise de formas derivadas destes verbos em documentos dos séculos XIII, XIV e XV, tanto de carácter literário como instrumental, levantadas a partir do *Corpus Informatizado do Português Medieval*, nesta comunicação tem-se como objetivos descrever como se distribuía essa distinção e explorar alguns dos processos de mudança linguística que intervieram na consolidação da distinção *ser/estar*. Esta investigação parte do quadro de estudos da gramaticalização para dar conta de alguns dos fenómenos observados no percurso destes verbos na língua portuguesa.

Como Veselinova (2006:122-126) nota, é um processo comum a diversas línguas verbos lexicais de posição estarem na origem dos verbos cópula, verificando-se que estes verbos começam a ser utilizados para denotar propriedades temporárias ou estados e, eventualmente, este uso é estendido para contextos que denotam propriedades mais permanentes. No caso das línguas românicas, em que os verbos *sedere* e *stare* estão na base dos verbos *ser*, estes verbos latinos teriam sofrido um processo de dessemantização e extensão, passando a significar genericamente 'estar', 'ficar' ou 'permanecer', o que os teria permitido aproximar-se de *esse*, convergindo com este em um único paradigma. Este quadro de estudos propicia, também, a análise do percurso evolutivo dos valores associados a estes verbos como resultado de um processo de competição entre formas e construções para a marcação dos mesmos valores, como sugerido por Brocardo (2014). No caso da língua portuguesa, a dessemantização dos verbos posturais *sedere*, *stare* e, também, *jazer* < *iacere* ('estar deitado') terá gerado uma sobreposição dos valores semânticos associados a estes verbos, o que eventualmente poderá ter determinado não só que

as formas menos frequentes se tenham tornado redundantes e caído em desuso – por exemplo, as formas de *sedere* que competiam com as *esse* (do género *sejo*, *siia*) –, mas também uma diferenciação dos valores associados aos paradigmas de *ser* e *estar*.

### Referências bibliográficas

- Brocardo, M. T. (2014). “Construções com *ser*, *estar*, *jazer* na história do português: notas em torno de inovação, persistência e obsolescência”. García, L.; Viñas, X. (eds.), *Língua, texto, diacronia. Estudos de linguística histórica. Revista Galega de Filoloxía*, monografía 9.
- Veselinova, L. (2006). *Suppletion in verb paradigms: Bits and pieces of the puzzle*. Typological Studies in Language, 67. Amsterdam: John Benjamins.

## Bare nouns and changes in the D system: insights from Old Portuguese

Clara Pinto

*Universidade Nova de Lisboa*

Bare nouns have been an ongoing topic of research ever since Abney's (1987) DP hypothesis. The existence of a DP projection has been consensually adopted in the literature, as well as the idea that “DPs can be arguments, NPs cannot”, following Longobardi (2001:581).

In Contemporary European Portuguese (CEP), bare singulars are generally ruled out from argument positions and bare plurals are only accepted under special circumstances. However, data from Old Portuguese suggest that bare singulars were more freely allowed. Examples (1) and (2) illustrate a contrast between Old and Contemporary Portuguese.

- (1) Emtã metou mão a espada.  
 So put.<sup>3sg</sup> hand to.the sword  
 So he grabbed the sword.

(DSG69)

- (2) [...] meteu a mão no bolso, retirou o pano branco  
 put.<sup>3sg</sup> the hand in.the pocket took.<sup>3sg</sup> the cloth white  
 [...] he put his hand on his pocket and took the white cloth.

(CRPC, L0393)

In both examples the common noun *mão* ‘hand’ is the internal argument of the verb *meter* ‘to put’. While in (1) the common noun *mão* appears under a bare form, in (2) it is preceded by the definite determiner *a* ‘the’. The occurrence of the common noun without the determiner would render the sentence ungrammatical, or, at least, marginal, as illustrated in (3).

- (3) #/\* meteu mão no bolso  
 put.<sup>3sg</sup> hand in.the pocket

The same situation is verified when the common noun is an external argument, as in (4) and (5). The use of the bare noun *abadesa* ‘abbess’ as the subject of the sentence is possible in Old Portuguese (4) but it would be ungrammatical in CEP, since the presence of a determiner (or other specifier) is mandatory, as in (5).

- (4) E abadesa choraua mujto com prazer tanto que uio Lançarot  
 and abbess cried.<sup>3sg</sup> much with pleasure such that saw.<sup>3sg</sup> Lancelot

And the abbess cried so much of joy as soon she saw Lancelot

(DSG04)

- (5) “Quando cheguei ao mosteiro a abadessa disse-me:”  
when arrived.<sup>1sg</sup> to.the monastery the abbess told.me.<sup>1sg.dat.</sup>  
When I arrived to the monastery the abbess told me:

(Corpus do Português)

In this presentation I will compare the frequency of bare nouns in syntactically annotated texts from 13<sup>th</sup> to the 18<sup>th</sup> century, extracted from diachronic corpora. The results point to a progressive decrease of singular bare nouns as arguments, suggesting a possible change from a system that allowed null Ds to one that demands this position to be lexically filled, as is the case of CEP.

## References

- Abney, S. (1987). *The English Noun Phrase in its Sentential Aspect*. PhD dissertation. MIT.
- Corpus do Português = Davies, M.; Ferreira, M. (2006-). *Corpus do Português*.
- DSG = Piel, J.-M.<sup>a</sup>; Nunes, I. (1988). *Demanda do Santo Graal*. Lisboa: INCM.
- Longobardi, G. (2001). “The Structure of DPs: some principles, parameters and problems”. Baltin, M.; Collins, C. (eds.), *The Handbook of Contemporary Syntactic Theory*. MA/Oxford: Blackwell, 562-603.

Aula Baratto

**Sesión Sociolingüística Histórica (1)**

17:00 – 18:30

## **¿Cuándo empieza el español de Cataluña? Un cotejo de corpus epistolares entre Benito Pérez Galdós, Joan Maragall y Narcís Oller**

Mar Garachana / Álvaro Octavio de Toledo y Huerta  
*Universitat de Barcelona / CSIC-ILLA*

El objetivo de la presente comunicación es analizar la lengua de dos escritores decimonónicos a fin de evaluar las similitudes y las diferencias existentes entre el español de regiones monolingües y el de territorios peninsulares en los que el español convive con otras lenguas. En esta comunicación, la comparación se circunscribirá al español y el catalán. La publicación de corpus epistolares extensos que ha venido realizándose a lo largo de los últimos años permite establecer este cotejo que, en este caso particular, se va a centrar, fundamentalmente, en analizar la correspondencia de dos escritores del XIX: el canario Benito Pérez Galdós y el catalán Joan Maragall.

El ochocientos marca la consolidación del castellano como dialecto terciario en Barcelona, de modo que, a partir de este momento, es este *primer español moderno* (Octavio de Toledo y Huerta 2016, 2019) el español deja de ser la lengua de una minoría selecta y se generaliza entre un conjunto relevante de la población (Kailuweit 1996, 1997). Así las cosas, un estudio comparativo resulta relevante, pues la lengua que reflejan las cartas escritas por Joan Maragall no solo permite caracterizar el

español en general, sino también medir la profundidad histórica de algunos rasgos que han sido considerados como definitorios del español de regiones en las que el castellano convive con el catalán. Más específicamente, con este trabajo aspiramos a determinar las diferencias que existían entre el español barcelonés y el de otras regiones monolingües peninsulares en lo relativo al marcado diferencial de objeto y al empleo de ciertas perífrasis verbales (la deóntica *haber de* + INF, la prospectiva *ir a* + INF y *deber (de)* + INF en sus usos epistémicos). Para llevar a cabo este trabajo, realizaremos un vaciado exhaustivo de la correspondencia publicada de Pérez Galdós y de Maragall (Smith et al. 2016; Garachana et al. 2022), que completaremos en puntos concretos con la mantenida entre el propio Galdós y otro escritor catalán, Narcís Oller (Shoemaker 1964).

### Referencias bibliográficas

- Garachana, M.; Artigas, E.; Cabré, L.; Bastardas, J. (2022). *Joan Maragall. Correspondencia*. Beihefte zur Zeitschrift für Katalanistik.
- Kailuweit, R. (1996). "El castellano de Barcelona en torno a 1880. La formación de un dialecto terciario". Alonso González, A. ; Castro Ramos, L.; Gutiérrez Rodilla, B.; Pascual Rodríguez, J. A. (eds.), *Actas del III Congreso Internacional de Historia de la Lengua Española (Salamanca, 22-27 de noviembre de 1993)*. Madrid: Arco Libros, 737–746.
- Kailuweit, R. (1997). *Vom EIGENEN SPRECHEN: Eine Geschichte der spanish-katalanischen Diglossie (1759-1859)*. Frankfurt am Main/New York: Lang.
- De Toledo, O. Huerta, Á. S. (2019). "Sintaxis de la prosa del instante: la lengua de una tradición efímera en los albores del siglo XIX". *Anuari de Filologia. Estudis de Lingüística*, 9, 91–144.
- Shoemaker, W. H. (1964). "Una amistad literaria La correspondencia epistolar entre Galdós y Narciso Oller". *Boletín de La Real Academia de Buenas Letras de Barcelona*, 30, 247-306.
- Smith, A. E.; Rodríguez Sánchez, M. Á.; Lomask, L. (2016). *Benito Pérez Galdós. Correspondencia*. Madrid: Cátedra.

---

### Aproximación al léxico de los papeles de la Inquisición: las relaciones de causas y alegaciones fiscales de tribunales andaluces como fuente de estudio

Pilar López Mora  
*Universidad de Málaga*

A partir de un corpus de alegaciones fiscales y relaciones de causas de los tribunales de Córdoba, Granada y Sevilla, del último periodo de la Inquisición (ss. XVIII y principios del XIX), estudiaremos voces y locuciones características de este tipo documental. Las expresiones propias, latinismos y denominaciones más frecuentes de los diversos casos (supersticiones, proposiciones heréticas, juicios a judaizantes y moriscos, bigamia, etc.) conforman un elenco de interés para la lexicología histórica y de su estudio se desprende un modo de expresión propio de una tradición textual con una identidad muy pronunciada y con características diferenciales que impone una profundización en su sentido particular y el entendimiento de la última etapa del Santo Oficio.



## Referencias bibliográficas

- Carrasco Hernández, I. (2010). "El discurso referido en una narración inquisitorial del siglo XVII". *AnMal Electrónica*, 29, 141-157.
- Eberenz, R.; De la Torre, M. (2003). *Conversaciones estrechamente vigiladas: Interacción coloquial y español oral en las actas inquisitoriales de los siglos XV a XVII*. Zaragoza: Pórtico.
- Gacto, E. (2012). "El procedimiento judicial en los Tribunales del Santo Oficio". Gacto, E., *Estudios jurídicos sobre la Inquisición Española*. Madrid: Dykinson, 167-182.
- Galende, J. C. (2001). "El proceso inquisitorial a través de su documentación. Estudio diplomático". *Espacio, tiempo y forma*, serie IV, Historia moderna, 14, 491-517.
- García-Muñoz, E. (2021). "Símbolos y signos ortográficos en la alegación fiscal del XVIII". *Scriptum digital*, 10, 27-52.
- Henningesen, G. (2014). "La inquisición y las brujas". *eHumanista: Journal of Iberian Studies*, 26, 133-152.
- Millán, J. M. (1983). "Crisis y decadencia de la Inquisición". *Cuadernos de investigación histórica*, 7, 5-18.
- Moreno, N. (1977). *Catálogo de Alegaciones fiscales*. Madrid: Servicio de Publicaciones del Ministerio de Educación y Ciencia.
- Pérez Martín, A. (1989). "La doctrina jurídica y el proceso inquisitorial". Escudero, J. A. (coord.), *Perfiles Jurídicos de la Inquisición española*. Madrid: Universidad Complutense, Instituto de Historia de la Inquisición, 279-343.
- Ramis, R. (2011). "Las alegaciones fiscales del Tribunal de la Inquisición de Mallorca". *Cuadernos de Historia del Derecho*, 18, 285-299.

---

## Cartas de particulares al Ayuntamiento de Málaga durante el siglo XIX

Livia C. García Aguiar  
*Universidad de Granada*

Los corpus de documentación municipal del s. XIX ofrecen grandes posibilidades para la investigación de los rasgos lingüísticos de clases sociales con distintos niveles de instrucción. Esto se debe a que, entre las colecciones de los archivos municipales, se pueden encontrar cartas que los ciudadanos enviaron al órgano rector con distintos objetivos: solicitar favores, certificar el correcto funcionamiento de las instituciones que gobernaban, informar de asuntos de variada naturaleza, etc.

El aumento de la alfabetización y el mayor acceso a la cultura durante el XIX se refleja también en la documentación municipal, en el sentido de que se produce un incremento en el número de cartas autógrafas o delegadas gráficamente, pero que incluyen la firma del emisor, lo que nos ofrece información sobre las habilidades de escritura de los distintos sexos y clases sociales que remitían sus cartas al Ayuntamiento. Así, estas cartas son igualmente valiosas para el estudio de la escritura femenina, cuyos escasos testimonios hasta este período quedan prácticamente relegados a los producidos por las mujeres de la nobleza. A partir de

los siglos XVIII y XIX, sin embargo, se multiplican los corpus epistolares privados femeninos, de gran interés para la investigación filológica, como se refleja en recientes investigaciones al respecto (Gómez Seibane 2017; Ramírez Luengo 2017; Martínez Pasamar 2020; Garrido 2021; López Mora, en prensa). También el espacio público será lugar propicio para documentar la voz de la mujer en este período (García Aguiar, en prensa).

El objetivo de este trabajo es, por un lado, presentar este subcorpus de cartas particulares dirigidas al Ayuntamiento malagueño durante la primera mitad del XIX, integrado en el corpus del proyecto CODEMA19 (UMA20-FEDERJA-123); por otro, mostrar los primeros resultados del análisis de los rasgos lingüísticos y escriturales, para lo que tendremos en consideración los factores sociales de sus emisores.

### Referencias bibliográficas

- García Aguiar, L. C. (en prensa). "Escritura femenina en el espacio público. Documentos autógrafos y delegados en el Archivo Municipal de Málaga durante el siglo XIX (1802-1832)".
- Garrido Martín, B. (2021). "Cartas de mujeres y recursos para la intensificación y la expresión afectiva en un corpus del siglo XVIII". *Hipogrifo*, 9(1), 1027-1048.
- Gómez Seibane, S. (2017). "Cartas escritas por mujeres vascas en la primera mitad del siglo XIX". Almeida, B.; Díaz Moreno, R.; Fernández López, M<sup>a</sup>. C. (eds.), *"Cansada tendré a vuestra excelencia con tan larga carta". Estudios sobre aprendizaje y práctica de la escritura por mujeres en el ámbito hispánico (1500-1900)*. Lugo: Axac, 133-149.
- López Mora, P. (en prensa). "Aproximación a la lengua de las cartas de sor Dolores Fernández de Córdoba y su familia (1759-1830)".
- Ramírez Luengo, J. L. (2017). "Mujeres y ortografía en la América decimonónica: la puntuación en las cartas de María Guadalupe Cuenca (Buenos Aires, 1811)". Almeida, B.; Díaz Moreno, R.; Fernández López, M<sup>a</sup>. C. (eds.), *"Cansada tendré a vuestra excelencia con tan larga carta". Estudios sobre aprendizaje y práctica de la escritura por mujeres en el ámbito hispánico (1500-1900)*. Lugo: Axac, 85-98.
- Martínez Pasamar, C. (2020). "Tuya, tuya aunque no quieran. Voz femenina e intensificación en correspondencia amorosa del español moderno". *ELiES*, 42, 87-111.

---

Aula Berengo  
**Sesión Diseño de Corpus (2)**  
17:00 – 18:30

### Colaboración y representatividad en el renovado Corpus CHARTA: los aportes de CODEA+, CORDINA y CORDEX

Víctor Caballero Gómez / Ricardo Pichel / Diego Sánchez Sierra  
*Universidad Complutense de Madrid / Universidad de Alcalá*

El Corpus Hispánico y Americano en la Red: Textos Antiguos (CHARTA: <https://www.corpuscharta.es/>), formado por las aportaciones de los más de veinte grupos que componen la Red Internacional CHARTA, cuenta hasta el momento con un total de 2076 documentos. De ellos, CHARTA ha ofrecido, en paralelo, una transcripción paleográfica y una presentación crítica, distintos estadios de un mismo proceso editorial cuyos principios metodológicos fueron consensuados por la mencionada Red CHARTA (2013). Los particulares intereses de cada uno de los grupos provocaron, no obstante, la sobrerrepresentación de determinados tipos de testimonios. Así, por ejemplo, los documentos procedentes de España ascendían al 95% y los datados en el siglo XIII se situaban en el 54%.

El proyecto «CHARTA 3.0: de la edición digital a la web semántica», dirigido por Ricardo Pichel desde la Universidad de Alcalá y financiado por la Comunidad de Madrid durante el período 01/01/2020–31/03/2022 (ref. CM/JIN/2019-008), del cual es resultado esta comunicación, tiene como objetivo principal la migración del Corpus CHARTA a la plataforma TEITOK (Janssen 2014); pero, además, se ha propuesto revisar la representatividad cronológica y geográfica de los testimonios accesibles a través de esta nueva vía de acceso. A la finalización del proyecto dispondremos, de este modo, de alrededor de 1300 documentos, de los cuales, los testimonios españoles descenderán del 95 al 77% y ninguno de los siglos representados (del IX al XIX) contará con más del 19% de documentos respecto del total. Ello supone, por ejemplo, un aumento de casi el 200% en el volumen de textos americanos incluidos en CHARTA. Esta actualización ha sido posible gracias a la colaboración, entre otros, del resto de los proyectos actualmente coordinados por el Grupo de Investigación de Textos para la Historia del Español (GITHE); esto es, «HERES: patrimonio textual ibérico y novohispano. Recuperación y memoria» (ref. 2018-T1/HUM-10230) y «Corpus de documentos españoles anteriores a 1900 (CODEA+ 2000)» (ref. FFI2017-82770-P), desde los cuales se desarrollan y amplían el Corpus Diacrónico del Español de Norteamérica (CORDINA: <https://corpora.uah.es/cordina/>) (Sánchez Sierra 2021: 152-153), el ya citado CODEA+ (<https://corpuscodea.es/>) y el Corpus Diacrónico de Extremadura (CORDEX).

Con estas actuaciones, pretendemos que CHARTA se aproxime a un corpus general en cuanto a su temática, restringido en relación con su magnitud, abierto atendiendo a su evolución y equivalente en lo relativo a su distribución geográfica y cronológica (Torruella Casañas 2017: 43-50). El propósito de nuestra comunicación es mostrar esta actualización operada sobre CHARTA, las nuevas posibilidades de investigación diacrónica y diatópica que ofrecerá, así como anticipar los retos que deberían abordarse para su progresiva mejora.

## Referencias bibliográficas

- Janssen, Maarten (2014): *TEITOK. A Tokenized TEI environment*, <http://www.teitok.org/>.
- Red CHARTA (2013<sup>2</sup> [2011]): *Criterios de edición de documentos hispánicos (Orígenes - siglo XIX)*, <https://www.redcharta.es/criterios-de-edicion/>.
- Sánchez Sierra, Diego (2021): «Aproximación a la configuración léxica del español en el suroeste de los Estados Unidos (1733-1900)», *Cuadernos del Instituto Historia de la Lengua* 14, pp. 147-185.
- Torruella Casañas, Joan (2017): *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística*. Frankfurt am Main: Peter Lang.

## Corpus y edición digital de textos antiguos gallego-portugueses. CGPA y EGPA: tradición y avances

Ricardo Pichel / Xavier Varela Barreiro

*Universidad de Alcalá / Universidade de Santiago de Compostela*

Durante las dos últimas décadas, la historia de la lengua latina y de las variedades romances (gallego-portugués y castellano) en el territorio de la *Gallaecia Magna* ha sido recogida, registrada y codificada en diferentes corpus lingüísticos para facilitar su estudio sistemático, en particular, el CODOLGA (*Corpus Documentale Latinum Gallaeciae*) y el TMILG (*Tesouro Medieval Informatizado da Lingua Galega*), este último ampliado, desde 2015, con los textos gallego-portugueses antiguos producidos en Portugal y actualizado en lo que hoy es el CGPA (*Corpus Galego-Portugués Antigo*).

En esta comunicación se presentará, por una parte, el estado actual de este último corpus multilingüístico, actualmente en proceso de convergencia con el CODOLGA en una misma plataforma. En su segunda parte, la exposición se centrará en la presentación de los últimos desarrollos en términos de corpus textual ecdóticamente multidimensional y de escritorio de multiedición digital proyectados a partir de las experiencias anteriores: el EGPA (*Escritorio Galego-Portugués Antigo*), anteriormente denominado COTAGAL (*Corpus de Textos Antigos de Galiza*), en el que se recoge la producción literaria y documental gallego-portuguesa producida en el espacio cultural histórico de Galicia y Portugal durante las edades Media y Moderna.

El objetivo es hacer balance de los últimos avances y del impacto de las mejoras alcanzadas en los entornos digitales de la codificación, visualización y explotación de este importante acervo textual, haciendo hincapié especialmente en las iniciativas colaborativas que han abierto un nuevo horizonte favorecedor del intercambio de experiencias y del reparto racional de responsabilidades y tareas en la implementación de nuevos protocolos de trabajo en la edición digital y en la minería de datos.

### Referencias bibliográficas

- Gago Jover, F.; Pueyo Mena, J. (2018). "El Old Spanish Textual Archive, diseño y desarrollo de un corpus de textos medievales: el corpus textual". *Cuadernos del Instituto de Historia de la Lengua*, 11, 165-209.
- Gago Jover, F.; Pueyo Mena, J. (2018). "El Old Spanish Textual Archive, diseño y desarrollo de un corpus de textos medievales: lematización y etiquetado gramatical". *Scriptum digital*, 7, pp. 25-35.
- Janssen, M. (2014). *TEITOK. A Tokenized TEI environment*. <http://www.teitok.org/>.
- Red CHARTA (2013<sup>2</sup> [2011]): *Criterios de edición de documentos hispánicos (Orígenes - siglo XIX)*. <https://www.redcharta.es/criterios-de-edicion/>.
- Torruella Casañas, J. (2017): *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística*. Frankfurt am Main: Peter Lang.
- Janssen, M.; Vaamonde, G. (2020). "Da edición dixital á análise lingüística. A creación de corpus históricos na plataforma TEITOK". Álvarez, R.; González Seoane, E. (eds.), *Calen barbas, falen cartas. A escrita en galego na Idade Moderna*.

## Lingüística de corpus con ojo crítico: el caso del fuero de Medinaceli

Mónica Castillo Lluch / Andrea Escobar Castillo  
*Université de Lausanne*

El objetivo fundamental de esta comunicación será presentar una nueva edición del fuero de Medinaceli concedido a la villa navarra de Murillo el Fruto. Este texto del siglo XIII se conserva en un manuscrito en el Archivo Real y General de Navarra (Cámara de Comptos Reales de Navarra, cajón 1, núm. 115) y de él existen dos ediciones: la primera realizada por Tomás Muñoz y Romero en su *Colección de fueros municipales y cartas pueblas de los reinos de Castilla, León, Corona de Aragón y Navarra* en el año 1847 (pp. 435-443), y la segunda publicada en 1965 por Ramón Menéndez Pidal y su equipo en la *Crestomatía del español medieval* (pp. 89-91). Por ser parcial la segunda (solo da 40 leyes de las 95), sigue siendo la primera, 175 años después de su publicación, la edición de referencia no solo para los historiadores del Derecho, sino también para la mayoría de los historiadores de la lengua, pues en el *Corpus diacrónico del español* (CORDE) de la RAE se volcó la versión del texto que ejecutó Muñoz y Romero.

La edición nos ha permitido, de una parte, conocer el proceso de edición de Muñoz y Romero y constatar sus numerosos defectos y, de otra, estudiar la lengua de ese texto una vez que lo hemos editado al completo con criterio filológico. Estos dos objetivos se añaden al principal de presentar nuestra nueva edición del fuero de Medinaceli concedido a Murillo el Fruto: de un lado, expondremos las deficiencias de la edición que hoy manejamos en el CORDE (incluidos los metadatos supuestamente relativos al texto) y, de otro, analizaremos las particularidades escritológicas del testimonio conservado y comentaremos los principales rasgos dialectales (navarros y vascos) que se registran en este manuscrito. En conclusión, subrayaremos la necesidad de revisar de modo crítico una parte importante de las ediciones que manejamos en los corpus diacrónicos informatizados y de reeditar esos textos con rigor filológico si pretendemos que sirvan de base sólida para nuestros estudios lingüísticos.

### Referencias bibliográficas

Muñoz y Romero, T. (1847). *Colección de fueros municipales y cartas pueblas de los reinos de Castilla, León, Corona de Aragón y Navarra*, t. I. Madrid: Imprenta de D. José María Alonso.

*Crestomatía* = Menéndez Pidal, R. (1965-1966): *Crestomatía del español medieval*, 2 vols., con la colaboración del Centro de Estudios Históricos, acabada y revisada por Rafael Lapesa y María Soledad de Andrés. Madrid: Gredos.

---

Jueves, 6 de octubre de 2022

Aula Baratto

**Sesión Sintaxis Histórica (2)**

9:00 – 10:30

## **The development of the absolute construction in (pre)classical Spanish: from a Latin calque towards a creative syntactic adaptation**

Marie Molenaers

*KU Leuven*

For my research project about the diachronic development of the absolute construction (AC), a non-finite subordinate clause with a certain level of syntactic independence from the main sentence, I composed a corpus of 15<sup>th</sup>-18<sup>th</sup>-century (pre)classical Spanish translations from Latin (Cicero's *De Officiis*, Seneca's *De Vita Beata*, *De Otio* and *De Providentia*, Boethius' *De Consolatione Philosophiae*). Given that the target texts are often of a non- or badly digitalized nature, I developed a *modus operandi* for the creation of documents suitable for linguistic analysis. The proposed presentation will give a detailed overview of this.

Translation studies are considered an ideal background to analyze linguistic changes (Baker 1993), but in contrast to the easy accessibility of the Latin source texts, several 15<sup>th</sup>-18<sup>th</sup>-century Spanish writings remain a challenge for scientific research. Previous analyses of ACs were restricted to one century (Pons Rodríguez 2006) or included only two works or authors (del Rey Quesada 2019). Furthermore, the Latin original was not or only partially examined, for example by merely focusing on the corresponding ACs between the source and target text (van de Pol 2016).

I compiled my diachronic corpus by transforming non-digitalized documents into digitalized ones with specific tools of the Adobe Scan Application. Afterwards, the OCR-editor functions of programs such as ABBYY FineReader and Transkribus enabled me to adjust unreadable letters and to manually delete glosses, comments and summaries, which optimize the accuracy of the search results and allow for the comparison between the different target texts. Moreover, the creation of editable files facilitates specific adjustments in the search mechanisms to precisely localize the ACs predicates, as well as the exact word counting, which is crucial for frequency calculations.

The digitalization makes a thorough scrutinization of barely investigated 15<sup>th</sup>-18<sup>th</sup>-century texts possible, which stimulates my diachronic investigation of ACs in various ways. In addition to previous AC studies, the focus is not only on the syntactic environment of translated ACs, but also on convergence or divergence with the original Latin AC. In the case of divergence, I found 'ex-novo' ACs without a Latin equivalent, whose evolution from typical Latinate participial towards gerundial predicates proves that Spanish ACs must thus no longer be considered as passively copied translational calques from Latin, but rather carved their own path of syntactic creativity.

Above this, my already digitalized documents are undoubtedly beneficial for researchers working on diachronic projects about other linguistic phenomena.

## References

- Baker, M. (1993). "Corpus Linguistics and Translation Studies". Baker M.; Francis G.; Tognini-Bonelli, E. (eds), *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 233-250.
- Del Rey Quesada, S. (2019). "Participial and gerundial clauses in sixteenth-century Spanish prose". *Belgian Journal of Linguistics*, 33, 43-81.
- Pons Rodríguez, L. (2006). "Una reflexión sobre el cambio lingüístico en el siglo XV". *Actas del V Congreso Andaluz de Lingüística General*, t. III. Granada: Lingvistica-Serie Collectae, 1563-1577.
- van de Pol, N. (2016). *The development of the Absolute Construction in English*. Leuven: KU Leuven. Faculteit Letteren.

---

## A bi-generational corpus investigation on factors constraining subject expression in European Portuguese spoken in Germany

Cristina Flores / Esther Rinke  
*Universidade do Minho / Goethe University*

The present study investigates subject expression in two generations of Portuguese migrants living in Hamburg, Germany. Based on a corpus of oral speech, we aim to assess whether second generation heritage speakers differ from first generation migrants with respect to the factors constraining subject realisation/omission in European Portuguese, a null subject language, in contact with German, a non-null subject language. The results do not reveal evidence in favour of ongoing language change, given that there are neither quantitative nor qualitative differences between the two generations of speakers. They show very similar overall rates of subject omission (around 67%) and they reveal sensitivity to the very same determining factors of subject pronoun realisation/ omission, namely person and number, verb type, switch reference (TC/TS) and distance. This finding is in line with previous corpus studies investigating the spontaneous speech of different generations of bilingual speakers or comparing monolingual and bilingual speakers of the same null subject language (e.g. Flores Ferrán 2004; Nagy 2014). We conclude that language contact per se does not necessarily lead to a diverging grammar at an inter-generational level, as long as stable input conditions allow for the acquisition of the constraints that are valid for null subject languages.

## References

- Flores-Ferrán, N. (2004). "Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact?". *Language Variation and Change*, 16, 49-73.
- Nagy, N. (2014). "A sociolinguistic view of null subjects and VOT in Toronto heritage languages". *Lingua*, 164B, 309-327.

## Existential structures in the Spanish of Mallorca then and now

Andrés Enrique-Arias / Silvia Perpiñán / Ileana Paul

*Universitat de les Illes Balears / Pompeu Fabra / Western University*

The Uniformitarian Principle claims, in very simple terms, that “the processes which we observe in the present can help us to gain knowledge about processes in the past” (Bergs 2012: 83). This principle has had ample application in historical linguistics; many language historians have adopted the view that language structure and the processes it undergoes are basically the same throughout the centuries. Labov’s famous formulation of the principle runs as follows: “the forces operating to produce linguistic change today are of the same kind and order of magnitude as those which operated five or ten thousand years ago” (Labov 1972: 275). This paper aims to explore the methodological opportunities and challenges of applying the Uniformitarian Principle in historical sociolinguistics and, more specifically, in the study of contact-induced variation in the past. The question to be addressed is basically the following: to what extent are the conditions leading to language transfer, and the effects resulting from it, stable across the centuries?

To this effect we use as case studies synchronic and historical analyses of the locative paradigm (Freeze 1991), which presents interesting microdifferences between Catalan and Spanish, two languages that are in a contact situation in Mallorca. In particular, the phenomena considered in Spanish in contact are: the use of the copula *ser* rather than *estar* to express location, as in (1); the combination of definite NPs with existential expressions with *haber* as opposed to *estar*, as in (2); and the use of the preposition *a* instead of *en* to express static location, as exemplified in (3).

(1) las sillas **eran** en casa del Marqués (1769 Felanitx)

the chairs were [=ser] in house of.the marquis

‘the chairs were at the marquis’s house’

(2) tú sabes desde la corte que es donde **hay la** gran política (1864 Palma)

you know from the Court which is where there.is the great politics

‘you know from the Court that that is where the great politics is’

(3) **estuvo a** Madrid un albañil de Binissalem (1861 Palma)

there-was at [=a] Madrid a mason from Binissalem

‘a mason from Binissalem was in Madrid’

These structures are extensively recorded in historical documents written by Mallorcans from different social classes, and also appear currently among Catalan-dominant speakers. Data for the comparative analysis comes from three newly developed corpora. For the historical analysis we use the Corpus Mallorca ([www.corpusmallorca.es](http://www.corpusmallorca.es)) which contains over a thousand texts in Spanish produced in the Catalan-speaking environment of Mallorca between 1640 and 1909. For the study of the current contact variety we extract data from the COSER corpus ([www.corpusrural.es](http://www.corpusrural.es)), a collection of sociolinguistic interviews with rural elderly informants, and from the PRESEEA-Palma corpus ([corpusmallorca.es/preseea](http://corpusmallorca.es/preseea)) with sociolinguistic interviews conducted in the city of Palma.

The quantitative analyses reveal that the degree of Catalan language dominance, formulated in terms of shorter periods of formal schooling and higher density of the Catalan-dominant population in the community where speakers live, is the main



predictor for the intensity of the transfer features. This effect is remarkably constant along the centuries and behaves similarly in both the historical and modern corpora; that is, some of the fundamental claims of modern (socio-)linguistics do seem to follow the Uniformitarian Principle, at least for this particular language contact situation. In sum, this paper demonstrates how a multidimensional approach to the study of linguistic variation that considers both synchronic and diachronic analytical techniques, allows us to identify and understand the structural, social, and cognitive mechanisms that influence processes of variation in language contact situations.

## References

- Bergs, A. (2012). "The Uniformitarian Principle and the Risk of Anachronisms". Hernández-Campoy, J. M. (ed.), *The Handbook of Historical Sociolinguistics*. Blackwell, 83-101.
- Freeze, R. (1992). "Existentials and Other Locatives". *Language*, 68(3), 553–595.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

---

Aula Berengo  
**Sesión Sintaxis Histórica (3)**  
11:00 – 12:30

### ***Sembrar la duda: estudio de una colocación en diacronía***

María Isabel Jiménez  
*Universidad Complutense de Madrid*

El español posee un conjunto de verbos, tales como *alegrar* y *asustar*, que sirven para expresar cambios de estado emocional o psicológico. Estos verbos se llaman "causativos emocionales", pues, como se ve en (1), la oración *El ruido asustó a Juan* puede parafrasearse por *El ruido causó que Juan sintiera miedo*.

- (1) a. El ruido asustó a Juan  
b. [El ruido] CAUSA [Juan siente miedo]

Sin embargo, no es la única opción que tiene el español para expresar la causatividad con términos emocionales o psicológicos. Los hablantes también pueden elegir una estructura formada por un verbo y un sustantivo, como en (2), que puede parafrasearse de la misma manera:

- (2) a. El ruido le dio miedo a Juan.  
b. [El ruido] CAUSA [Juan siente miedo]

Estas estructuras son bien conocidas ya en la bibliografía lingüística donde han recibido el nombre de "colocaciones verbo-nominales", definidas como "combinaciones restringidas de voces cuya frecuencia de coaparición es muy elevada" (NGLE 2009:2614).

En el español actual, un verbo altamente productivo en colocaciones de este tipo es *sembrar*. Lo encontramos unido a sustantivos como *terror*, *pánico*, *duda*, *dolor*,

*preocupación, desconfianza, sospecha, odio, dudas, confusión, inquietud, incertidumbre, desconcierto, paz, certeza, fe, etc (3):*

(3) a. colgar en las paredes de las oficinas unos cartelitos con máximas y recomendaciones breves, pero concluyentes, que tienden a *sembrar el pánico* entre las masas y entre las mesas. (Fernando Luque, *La lucha y el destino*, 1927)

b. No tardaron los apetitos en levantar voces de protesta y en sembrar la duda entre los ignorantes y los incapaces de toda comprensión redentora. (Rodrigo Soriano, *Un ensayo revolucionario*, 1932)

Este verbo nos resulta interesante por diferentes motivos. En primer lugar, por su alta frecuencia de uso en colocaciones (solo en el s.XX, más del 25% de todos sus ejemplos en un estudio de corpus son colocaciones). Y, por otra, por su valor metafórico (nos parece relevante que conforme construcciones mayoritariamente de valor negativo, siendo un verbo cuyo sentido literal originario tiene connotaciones positivas).

Los datos anteriores nos han llevado a desarrollar un estudio en diacronía de las colocaciones en las que participa este verbo de origen agrícola con sustantivos de polaridad negativa -tomando el sustantivo *duda* como muestra- desde la lengua latina hasta la actualidad con el objetivo, en primer lugar, de dilucidar los motivos que le llevaron a participar de estas construcciones y desde cuándo, y, por otro, cuáles son los nexos metafóricos que han hecho de él uno de los verbos preferidos para expresar relaciones de causación emocional o psicológica en español.

### Referencias bibliográficas

Real Academia Española; Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española. Morfología. Sintaxis*, 2 vols. Madrid: Espasa.

---

## La lexicalización de la deixis en los verbos de movimiento causado

Julio Torres Soler  
*Universidad de Alicante*

En la actualidad, los verbos *llevar* y *traer* en español son *verbos deícticos de movimiento*, puesto que incorporan información sobre las condiciones deícticas en las que se produce el desplazamiento al que hacen referencia. Las condiciones deícticas que gobiernan el uso de los verbos deícticos de movimiento en español no han sido las mismas en las diferentes etapas de la historia del idioma. Según Ricca (1993, cap. 5), en latín los verbos *ire* y *venire* muestran una oposición basada fundamentalmente en diferencias aspectuales, mientras que en las lenguas románicas se encuentra una tendencia a codificar una oposición basada cada vez más en la situación deíctica del momento de enunciación. En español clásico, si bien la oposición entre los verbos *ir* y *venir* ya era fundamentalmente deíctica, *venir* se utilizaba para designar desplazamientos hacia metas más diversas que en español moderno, incluyendo desplazamientos al lugar que ocupa la segunda persona en el momento de

enunciación. Esta tendencia se encuentra también en otras lenguas románicas de la península ibérica, como demuestran Montserrat Buendia (2005) en el catalán y, por otro lado, Moret Oliver e Ibarretxe-Antuñano (2018) en el aragonés.

En el presente trabajo, vamos a describir la lexicalización de la deixis en los verbos *llevar* y *traer* y su evolución en la historia del español. Para ello, llevamos a cabo el análisis cualitativo y cuantitativo de un corpus lingüístico formado por más de 2400 ocurrencias del verbo *llevar* y otras tantas de *traer* de todas las etapas del español. Los datos han sido extraídos del Corpus del Diccionario Histórico del Español y se limitan al ámbito geográfico de España. Nuestros resultados preliminares sugieren que *llevar* y *traer* experimentaron cambios semánticos por los que su oposición pasó de ser principalmente aspectual a ser deíctica. Estos cambios se explican por el diferente grado de conciencia y control que el conceptualizador tiene sobre la meta del desplazamiento: mayor en el caso del movimiento centrípeto y menor en el caso del movimiento centrífugo.

### Referencias bibliográficas

- Montserrat Buendia, S. (2005). "Entre la dixi i la definitud: els verbs de moviment resultatiu en català". *Caplletra*, 39, 61-83.
- Moret Oliver, M. T.; Ibarretxe-Antuñano, I. (2018). "Una aproximación desde la tipología semántica a la deixis (ir y venir) en aragonés y catalán medieval". *Actas del X Congreso Internacional de Historia de la Lengua Española (vol. 2)*. Zaragoza: Institución Fernando el Católico.
- Ricca, D. (1993). *I verbi deittici di movimento in Europa: una ricerca interlingüística*. Florencia: La nuova Italia.

---

## Variación histórica en las verbalizaciones: ¿qué aporta el prefijo al verbo?

Marina Espejel  
*Universitat de Girona*

El uso de corpus lingüísticos resulta imprescindible para el estudio de la variación histórica, juntamente con los enfoques teóricos y los estudios descriptivos, tanto sincrónicos como diacrónicos. La interpretación de datos nos permite, no solo trazar cuál ha sido la evolución de un fenómeno lingüístico, sino también, por ejemplo, los cambios tipológicos que ha habido en el sistema.

En esta comunicación se tratan las ventajas y desventajas que nos ofrecen los corpus de la RAE, de donde se extraen los datos diacrónicos (CDH, CORDE), así como el CE, que se ha utilizado para la lematización. Para ello, se examinan los datos extraídos en el estudio de la variación histórica que se produce en la rivalidad prefijal de las verbalizaciones.

El sistema lingüístico del español puede indicar cambio de estado a partir de distintos patrones morfológicos (*endurecer*, *engordar*, *acortar*, *clarear*, *sutilizar*). Además, una misma base puede presentar alternancia entre diferentes sufijos y prefijos (*entristecer* ~ *entristar* ~ *atristar*, cf. Malkiel 1941). Si se observa la variación a lo largo de la historia del español, se advierte que en los verbos corradicales que se generan en etapas

distintas, se produce una especialización de significado (cf. Pujol Payet 2019). En los casos que se hayan formado corradicales en la misma etapa, solo se mantienen si pueden expresar significados distintos, aunque puedan coincidir en algún aspecto: *aclarar* ~ *clarear*.

En esta investigación se propone que los verbos deadjetivales tienen una única interpretación semántica: cambio de estado (*engordecer* ~ *engordar*). Es este hecho el que fomenta que los afijos se desemanticen (cf. Iacobini 2021) y, en consecuencia, se produzca una mayor vacilación y polimorfismo en las primeras etapas del español. En los denominales, sin embargo, se defiende que los prefijos otorgan contenido léxico, puesto que el sustantivo es poliédrico (*empedrar* ~ *apedrar*). Asimismo, se explica la variación que se advierte sincrónicamente, ya que un enfoque diacrónico permite observar la productividad de cada uno de los patrones, así como mostrar su periodización, reduciendo la aparente redundancia que se produce.

### Referencias bibliográficas

- CDH = Real Academia Española. *Corpus del Nuevo Diccionario Histórico del Español*.  
<http://www.rae.es/recursos/banco-de-datos/cdh>.
- CE = Davies, M. *Corpus del español*. <http://www.corpusdelespañol.org>.
- CORDE = Real Academia Española. *Corpus Diacrónico del Español*.  
<http://www.rae.es/recursos/banco-de-datos/corde>.
- Iacobini, C. (2021). "Parasynthesis in Morphology". Lieber, R.; Fábregas, A.; Gagné C. L.; Masini, F.; Arndt-Lappe, S. (eds.), *The Oxford Encyclopedia of Morphology*, Oxford: Oxford University Press, vol. 1, 765-779.
- Malkiel, Y. (1941). "Atristar – entristecer: adjectival verbs in Spanish, Portuguese and Catalan". *Studies in Philology*, 38, 429-461.
- Pujol Payet, I. (2019). "Sobre los verbos complejos denominales: ¿una clase verbal?" Adelstein, A.; Bernal, E.; Sinner, C. (eds.), *Clases y categorías en la formación de palabras en español*. Leipzig: Leipziger Universitätsverlag.

---

Aula Baratto  
**Sesión Sociolingüística Histórica (2)**  
11:00 – 12:30

### **Estudio socioestilístico sobre el flamenco: fenómenos fonéticos para explicar la evolución de los sonidos del género**

Elena Fernández de Molina Ortés  
*Universidad de Granada*

Desde los orígenes del cante flamenco (tratado como género actual) en el siglo XVIII, han sido abundantes y variados los trabajos que han analizado esta forma de expresión (antropológicos, culturales, sociales etc.). Desde un punto de vista lingüístico, sin embargo, no son muchos los trabajos sobre el género y, los que han sido publicados, se relacionan más con aspectos léxicos y la influencia del caló en el español. Sin embargo, y desde nuestro punto de vista, uno de los rasgos más representativos del flamenco es su pronunciación; los cantaores no son solo conocidos por su forma de cantar, sino también por los sonidos que usan, que los

ubican en un espacio geolingüístico concreto (el español meridional, sobre todo el de Andalucía).

En esta investigación se propone realizar un estudio lingüístico con el que 1) poder observar qué fenómenos fonéticos se asociaban con el flamenco en los orígenes del canto (siglo XVIII) y 2) comprobar si las variantes del género se conservan en la actualidad. En este caso, y teniendo en cuenta la zona geográfica en la que se origina el flamenco, H1) los sonidos que se usen serán propios del área occidental de Andalucía (aspiraciones, seseo y ceceo, fricativización de /tʃ/). H2) Asimismo, un género tradicional como el flamenco ha debido conservarse de una forma muy similar al de los orígenes; de hecho, en las primeras épocas, de forma inconsciente, los cantaores tradicionales que llevaron el flamenco desde Andalucía a otras zonas posiblemente convirtieron ciertos fenómenos fonéticos en rasgos estilísticos del canto. H3) Sin embargo, aunque el flamenco tradicional pervive en la actualidad, convive, también, con otros tipos de flamenco (flamenco fusión, flamenco tune), en donde parece que se conservan ritmos del género pero cambian el estilo, los contenidos y el tipo de música. En estos nuevos estilos del flamenco, seguramente, se observarán menos fenómenos del canto tradicional aunque, al igual que se conservan ritmos del género, también se podrán escuchar sonidos tradicionales.

Para llevar a cabo estos objetivos y comprobar las hipótesis de investigación, nos serviremos de la metodología propuesta por la sociolingüística de la tercera ola, mediante la cual analizaremos cómo ha variado un estilo musical, el flamenco, a lo largo del tiempo. Se incluyen aquí variables no solo lingüísticas y sociales, sino también antropológicas, para estudiar un género musical (Cutler 1999b, 1999a; Gibson, Bell 2012; Jansen, Westphal 2017; Maeve Eberhardt 2020; Samy Alim 2002; Simpson 1999; Trudgill 1980, 1983) Concretamente, y para hacerlo, hemos transcrito y analizado un corpus sonoro compuesto por más de 600 horas de grabación. Para hacerlo, se han establecido dos tipos de flamenco (flamenco tradicional y flamenco nuevo -relacionado, en este caso, con el flamenco fusión, generado a partir de los años 70) y, además, se han distribuido 3 generaciones: la tercera, la compuesta por cantaores nacidos en el siglo XIX y el siglo XX, que serían los impulsores de la estilización del canto; la segunda, integrada por sus herederos, los cantaores nacidos a partir de los años 40 y 50 y la primera generación, que está constituida por cantaores menores de 40 años en la actualidad.

El análisis de este corpus revela que, efectivamente, algunas variantes usadas en el flamenco de los orígenes han pervivido durante más de dos siglos, hasta la actualidad (ver Fernández de Molina Ortés 2020, 2022). De hecho, se advierten tendencias muy interesantes entre las distintas generaciones y procedencias de los cantaores: por ejemplo, existen fenómenos con una gran extensión en los orígenes que, sin embargo, en la actualidad se usan con menos frecuencia y, también, nuevos sonidos que aparecen en Andalucía (fricativización de /s/ implosiva, por ejemplo) y que se convierten, poco a poco, en representativos del género.

### Referencias bibliográficas

- Cutler, C. A. (1999a). "Yorkville Crossing: white teens, hip hop, and African American English". *Journal of Sociolinguistics*, 3(4), 428-442.
- Cutler, C. A. (1999b). "Yorkville crossing: White teens, hip hop and African American English". *Journal of Sociolinguistics*, 3, 428-441.
- Fernández de Molina Ortés, E. (2020). "Los sonidos del flamenco: análisis fonético de "los orígenes" del canto". *Cultura, Lenguaje y Representación*, 24, 53-74.

- Fernández de Molina Ortés, E. (2022). "El flamenco puro desde una perspectiva lingüística. Fenómenos fonéticos como rasgos de identidad". Popek-Bernat K. (Ed.), *Normatividad, equivalencia y calidad en la traducción e interpretación de lenguas ibéricas*. Frankfurt am Main: Peter Lang, 54-75.
- Gibson, A., Bell, A. (2012). "Popular music singing as referee design". Hernández-Campoy, J.M.; Cutillas-Espinosa, J.A. (eds.), *Style-Shifting in Public: New Perspectives on Stylistic Variation*. Amsterdam/Philadelphia: John Benjamins, 139-164.
- Jansen, L.; Westphal, M. (2017). "Rihanna Works Her Multivocal Pop Persona: A Morpho-syntactic and Accent Analysis of Rihanna's Singing Style". *English Today*, 33(2), 46-55.
- Maeve Eberhardt, M. V.-M. (2020). "I ain't sorry": African American English as a strategic resource in Beyoncé's performative persona. *Language & Communication*, 77, 68-78.
- Samy Alim, H. (2002). "Street-conscious copula variation in the hip hop nation". *American Speech*, 77(3), 288-304.
- Simpson, P. (1999). "Language, culture and identity: With (another) look at accents in pop and rock singing". *Multilingua*, 18(4), 343-367.
- Trudgill, P. (1980). "Acts of conflicting identity: a sociolinguistic look at British pop songs". Sugathapala de Silva, M. V. (ed.), *Aspects of Linguistic Behaviour: Festschrift for R.B. Le Page. York Papers in Linguistics*, 9. York: University of York.
- Trudgill, P. (1983). "Acts of Conflicting Identity: The Sociolinguistics of British Pop-song Pronunciation". Trudgill, P. (ed.), *On Dialect: Social and Geographical Perspectives*. New York: Basil Blackwell y New York University Press, 141-160.

---

### **Coloquialización y cambio lingüístico en el siglo XX: primer acercamiento al caso de las narraciones futbolísticas en español peninsular**

Shima Salameh Jiménez  
*Universitat Jaume I*

La reciente incorporación del enfoque pragmático y sociolingüístico a la investigación diacrónica ha puesto de manifiesto la importancia del tratamiento del cambio lingüístico a través de parámetros de variación en diferentes niveles (geográfico, registro, género discursivo, nivel sociocultural, etc.). Uno de los procesos que aún a cambio lingüístico, registro y género discursivo es la *coloquialización*, definida como "la incorporación de características propias del habla en el lenguaje escrito" en ensayos, diarios, prosa, ficción, etc. (Traugott 2010), o, en un sentido más amplio, "la incorporación de rasgos coloquiales en situaciones comunicativas habitualmente formales" (Briz 2010).

Algunos trabajos apuntan a una coloquialización del español peninsular en el último cuarto del siglo XX, muy relacionada con la consolidación de los medios de comunicación de masas, convertidos en herramienta de innovación y difusión del cambio (Pons 2014: 1002, 1003). Basta con comparar prensa, radio y televisión actuales y antiguas para extraer diferencias en el uso de la lengua que van de "una estética y verbalización lingüística muy formales" a una inmediatez comunicativa en

diferentes géneros discursivos y formatos (López Serena 2014: 39) que supone una importante transformación, dado que los medios, con estos nuevos rasgos, se convierten en modelos comunicativos para espectadores (nuevos saludos, despedidas, fórmulas de tratamiento, difusión de léxico, etc.). Por tanto, su coloquialización permite delimitar el proceso general en el español peninsular y obtener un mapa evolutivo más completo.

En esta comunicación abordaremos la coloquialización a través de las retransmisiones futbolísticas: aunque el vocabulario del fútbol en español se ha trabajado en detalle (Nomdedeu 2019), su tratamiento ha sido principalmente léxico, sin un estudio diacrónico y pragmático de las producciones escritas o audiovisuales. Las retransmisiones (narraciones del partido) constituyen una fuente de usos orales reales, asociados a los medios de comunicación de masas, que reflejan un cambio formal > informal en las últimas décadas, y son un motor de difusión a raíz de su impacto en la sociedad. Hemos recopilado cerca de 30 horas de narraciones audiovisuales (TVE, A3, Canal +, Cope, RNE, etc.) desde 1950 para un primer análisis en tres niveles: prosodia, léxico y pragmática. En líneas generales, hay coloquialización en diferentes aspectos: ritmo más dinámico y veloz en la narración, uso de léxico y expresiones informales, más modalización y valoración, abandono del ustedeo, entre otros.

### Referencias bibliográficas

- Briz, A. (2010). "Lo coloquial y lo formal, el eje de la variedad lingüística". Castañer Martín R. M.<sup>a</sup>; Lagüéns Gracia, V. (coords.): *De moneda nunca usada: Estudios dedicados a José M<sup>a</sup> Enguita Utrilla*. Zaragoza: Instituto Fernando El Católico/CSIC, 125-133.
- López Serena, A. (2014). "De la oralidad fingida a la oralidad simuladora de realidad: reflexiones en torno a la coloquialización del discurso como estrategia mediática". *Español actual: Revista de Español vivo*, 102, 37-75.
- Nomdedeu, A. (2019). "Las aportaciones de Antonio Viada (1902) a la estrategia españolizadora del lenguaje deportivo anglosajón. Estudio a propósito del diccionario histórico de términos del fútbol". *BRAE*, 723-770.
- Pons, S. (2014): "El siglo XX como diacronía: intuición y comprobación en el caso de *o sea*". *Rilce*, 30, 3, 985-1016.
- Traugott, E.Cl. (2010). "Grammaticalization". Jucker, A.; Taavitsainen, I. (eds.), *Historical Pragmatics*, 8, 97-126.

---

### El papel del individuo ante el cambio lingüístico: perfiles idiolectales en la correspondencia privada de seis personalidades coetáneas del siglo XIX

José Luis Blas Arroyo  
*Universitat Jaume I*

A partir de un corpus compuesto por cartas privadas escritas por seis personalidades coetáneas de la vida social y cultural española de la segunda mitad del siglo XIX, en el presente estudio se analizan las distribuciones idiolectales correspondientes a dos variables sintácticas sometidas a variación y cambio lingüístico en la época, aunque con diferencias relevantes en cuanto a su difusión. Mediante la comparación con las

tendencias centrales en la sociedad en ese periodo, obtenidas en estudios variacionistas previos, los objetivos del estudio se detienen específicamente en:

- a) Comparar los perfiles de variación idiolectales respecto a esas tendencias en dos momentos diferentes de la vida de los individuos (la etapa de supuesta consolidación de los repertorios comunicativos, al final de su juventud, y el periodo de redacción de las cartas, que, en todos los casos, se produjo en edades más avanzadas),
- b) Comprobar si estos perfiles coinciden en las dos variables estudiadas, o si, por el contrario, se advierten diferencias entre ambas,
- c) Analizar el grado de estabilidad o dinamismo de las distribuciones individuales en diferentes periodos de la producción epistolar.

Los resultados del estudio confirman la existencia de importantes diferencias en las distribuciones de las dos variables estudiadas. En relación con la primera —el uso de la variante con artículo en las subordinadas relativas oblicuas ('es la casa en *la* que nació'), frente a la variante tradicional, sin artículo ('es la casa en *que* nació')—, la mayoría exhibe un comportamiento claramente *refractario*, con frecuencias de uso prácticamente anecdóticas, que se sitúan incluso por debajo de las comunes en la época. Aun así, no faltan quienes se colocan en la vanguardia de este cambio lingüístico (*adeptos*), con cifras que superan claramente las comunes en cualquiera de los momentos que tomemos como referencia.

Las cosas son diferentes a propósito de los usos de la perífrasis *haber de + infinitivo*, tradicional en la historia del español, pero sometida ya en este periodo a considerable pugna por parte de *tener que + infinitivo*. Esta vez los perfiles mayoritarios se distribuyen entre los que llamamos *contemporizadores*, por el seguimiento que hacen de las distribuciones mayoritarias en la época, y *adeptos*, por sumarse con entusiasmo a las tendencias de cambio en el paradigma de las perífrasis. Con todo, en este caso se advierten diferencias en relación con el periodo tomado como referencia (18 años o una etapa posterior), lo que sugiere que no siempre las distribuciones se consolidan al final de la juventud, sino que pueden modificarse en etapas posteriores.

Al comparar los perfiles idiolectales ante los dos fenómenos de variación, se comprueba también que los hablantes no reaccionan de la misma manera ante ambos, de tal manera que un comportamiento determinado respecto a uno, no es incompatible con uno distinto ante otro. Finalmente, del análisis longitudinal de la correspondencia se desprende la convivencia de distribuciones que se mantienen básicamente inalteradas en el eje diacrónico, junto con otras en las que se aprecian cambios entre unos momentos y otros, generalmente en paralelo a las tendencias de cambio imperantes en la sociedad.

---

Aula Berengo

**Sesión Dialectología Histórica (1)**

11:00 – 12:30

**De *trinchar*, *trincher*. Algunos derivados en *ero/era* en documentación notarial seiscentista**

Alejandro Junquera Martínez



### Universidad de León

El *Corpus Léxico de Inventarios (CorLexIn)* se ha consolidado en los últimos años como una herramienta útil a la hora de atestiguar voces infrarrepresentadas —o directamente ausentes— en fuentes de carácter más general, hecho derivado, entre otros, de la tipología textual de la que parte (la prosa notarial, inventarios de bienes, cuyo registro se encuentra mucho más próximo a la realidad lingüística del periodo seiscentista). Asimismo, otra de las características que lo definen es su marcado carácter diatópico, fruto de la doble cala que efectúa a ambos lados del Atlántico, lo que también lo convierte en un recurso interesante para el estudio diatópico del léxico (Morala 2014; Morala, Egido 2018).

El objetivo que persigue la presente comunicación es ofrecer una caracterización de un corpus de voces (sustantivos y adjetivos) derivadas en *-ero/a* atestiguadas en el siglo XVII y escasamente documentadas en corpus generales (CORDE, CDH) a partir de tres perspectivas complementarias: lexicográfica, documental y, en aquellos casos en los que resulte pertinente, geográfico (ámbito geográfico restringido, uso diatópicamente marcado, fenómenos de variación formal, etc.).

En lo que respecta a la caracterización del elemento sufijal, además de su uso mayoritario junto a *-dero/a* a la hora de formar nombres de profesiones, la *NGLE* (§ 6.8i) indica que *-ero* y su forma femenina *-era* también son frecuentes en la formación —entre otros— de nombres recipiente, utensilios e instrumentos o prendas de vestir; de ahí que podamos encontrarnos con *cubiteras*, *sonajeros* y *hombreras*. Por otro lado (§ 7.8i y ss.), también pueden encontrarse ejemplos en los que *-ero/a* se emplea para generar adjetivos denominales, especialmente adjetivos de tipo relacional (barco *pesquero*, industria *maderera*, etc.) u otros ejemplos menos comunes como *besugero*, *trapero*, etc. (Perdiguero Villarreal 2012).

### Referencias bibliográficas

- CDH = Instituto de Investigación Rafael Lapesa de la Real Academia Española (2013)- *Corpus del Nuevo diccionario histórico* (CDH). <http://web.frl.es/CNDHE>.
- CORDE = Real Academia Española. *Corpus Diacrónico del Español* (CORDE). Disponible en: <http://corpus.rae.es/cordenet.html>.
- Morala Rodríguez, J.R. (2014). “El *CorLexIn*, un corpus para el estudio del léxico histórico y dialectal del Siglo de Oro”. *Scriptum digital*, 3, 5-28.
- Morala Rodríguez, J.R. (dir). *Corpus Léxico de Inventarios (CorLexIn)*. <https://apps2.rae.es/CORLEXIN.html>.
- Morala Rodríguez, J.R.; Egido Fernández, M.C. (2018). “El proyecto *CorLexIn* y la variación diatópica en el léxico del Siglo de Oro”. Corbella Díaz, M.D.; Fajardo Aguirre, A.; Langenbacher-Liebgoth, J. (eds.), *Historia del léxico español y humanidades digitales*. Bruselas: Peter Lang, 394-417.
- NGLE* = Real Academia Española; Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española* (2 vol.). Madrid: Espasa.
- Perdiguero Villarreal, H. (2012). “Palabras en *-ero/-era* en protocolos notariales de Castilla”. Campos Souto, M.; Mariño, R.; Pérez Pascual, J.I. et al (eds.), “*Assí como es de suso dicho*”: *Estudios de morfología y léxico en homenaje a Jesús Pena*. San Millán de la Cogolla: Cilengua, 381-391.

## Changes in Null Subjects in Latin American Spanish

Gemma McCarley  
*Universität Konstanz*

The phenomenon of partial null subject languages that increasingly allow both overt and null subject pronouns in the same contexts has been observed in varieties such as Dominican Spanish (Toribio 2000). Given Latin American Spanish's rich history of language contact over the last five centuries, the observed reduction in use of null subjects is likely a result of contact-induced simplification. Following Trudgill's (2011) sociolinguistic typology, this contact would be characterized as short-term adult second-language acquisition which is consistent with the scenario for the African slaves brought over to the Americas by the Spanish from the 16th century onward. Indeed, there are numerous non-standard varieties of Spanish spoken by Afro-Hispanic communities throughout Latin America. These Afro-Hispanic Languages of the Americas (AHLAs) have been described as the result of "conventionalized advanced second languages," much in line with Trudgill's proposal (Sessarego 2013, 2017). The L2-difficulty of the rules surrounding null subject use would have led adult learners to overproduce overt subject pronouns. Their errors would have then become nativized in the speech of the next generation. What is needed then is a robust diachronic analysis of null subjects. According to the literature we also expect a correlation between inversion and null subjects. The present study has constructed a corpus of 57 texts from eight countries over four centuries to evaluate the diachronic rate of null subjects and inversion in Latin American Spanish. The aim is to determine whether the corpus supports the hypothesis that the rate of overt subject pronouns (as well as SV order) has risen significantly since contact with African L2-speakers. The corpus is supplemented by transcriptions of previous fieldwork in current AHLA varieties as well as historical texts written in AHLA vernacular. Figures 1-2<sup>1</sup> show the distribution of referential null/overt subjects and SV/VS word orders from the Dominican Republic and Bolivia. This preliminary data has also shown an unexpected and very intriguing correlation between inversion rates and genre type: SV order is much more likely in non-literary documents (particularly legal court documents that summarize specific cases and witness testimony). Broadly, we expect to see the highest variation in the form of more overt subject pronouns and pre-verbal subjects in varieties with the highest Afro-Hispanic populations. This should form a spectrum from Caribbean to South American to Peninsular Spanish.

---

<sup>1</sup> The docID corresponds to the country, century, and title of a text (e.g. dr16ent = Dominican Republic, 16<sup>th</sup> century, Entremés). The document genre is bolded, and the golden bolded text is a transcript from Afro-Bolivian vernacular.

Figure 1: Null vs. Overt Subject Pronouns

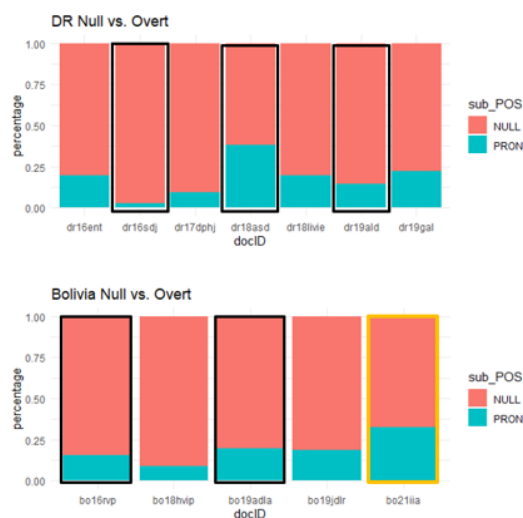
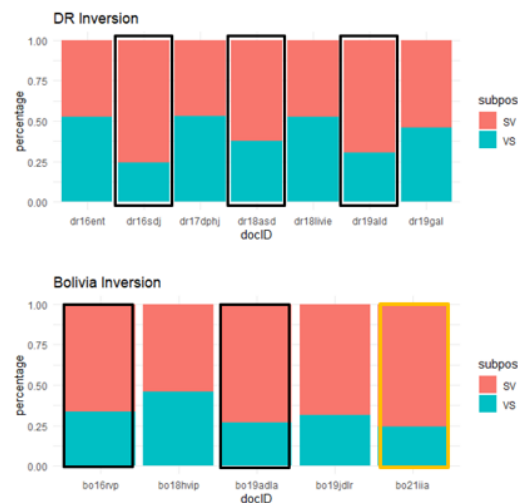


Figure 2: Inversion Rates



## References

- Sessarego, S. (2013). "Afro-Hispanic Contact Varieties as Conventionalized Advanced Second Languages". *IBERIA*, 5(1), 99-125.
- Sessarego, S. (2017). "Revisiting the Null Subject Parameter: New Insights from Afro-Peruvian Spanish". *Isogloss*, 3(1), 43-68.
- Toribio, A. J. (2000). "Setting parametric limits on dialectal variation in Spanish". *Lingua: International Review of General Linguistics*, 110(5), 315-341.
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: OUP.

## La terminología comercial en español a partir del corpus DIACOM-es: representación digital de variantes diacrónicas y diatópicas

Piero Costa León / Dunia Hourani / Matteo De Beni / Elisa Sartor  
*Università degli Studi di Verona*

Nuestra intención en esta comunicación es dar cuenta de los avances realizados en el proyecto DIACOM. Esta investigación, desarrollada en la Universidad de Verona en el marco de un "Proyecto de Excelencia" (2018-2022) financiado por el Ministerio italiano de Educación, Investigación y Universidades, está destinada a la compilación de dos corpus para el estudio y la representación del léxico especializado del comercio, uno en español y uno en francés. En concreto, nos centraremos, por un lado, en presentar los datos actuales del corpus de español (*DIACOM-es*), que incorpora textos de ámbito comercial de diversa tipología, redactados en los diferentes países hispanohablantes en Europa y América en tres cortes temporales: 1850-1914, 1945-1970 y 1990-2018.

Por otro lado, expondremos el modelo de representación gráfica de los términos que se ha diseñado en el marco del proyecto. En primer lugar, presentaremos la ficha terminológica proyectada para recoger la variación diacrónica y diatópica. Seguidamente, efectuaremos el análisis de una muestra de unidades terminológicas en las que se haya verificado tanto variación debida a la evolución conceptual en el curso del tiempo como motivada por razones geográficas. Por último, mostraremos cómo se ha registrado dicha información en la ficha terminológica y cómo los datos incorporados en la misma sirven para la posterior visualización digital de los términos en forma de red léxico-semántica.

### Referencias bibliográficas

- De Beni, M.; Hourani-Martín, D. (2020). “El discurso especializado del comercio en español: diseño de un corpus para su estudio diacrónico”. *Lexis. Revista de Lingüística y Literatura*, 44(1), 75-112.
- Frassi, P. (2021). “DIACOM-FR, une base de données terminologiques de type diachronique”. *Cahiers de Lexicologie*, 108(2), 23-49.
- Hourani-Martín, D.; Sartor, E. (2021): “Textos comerciales de Cuba y Colombia del siglo XIX en el corpus DIACOM: estudio de casos de variación terminológica”. De Beni, M.; Hourani-Martín, D. (eds.), *Corpus y estudio diacrónico del discurso especializado en español*. Frankfurt am Main: Peter Lang (Col. *Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation*, 161, 167-181.

---

Aula Baratto  
**Sesión Diseño de Corpus (3)**  
15:00 – 16:30

### La tecnología al servicio de la *collatio codicum*: el caso del Fuero Juzgo

Charles Mabile  
*Université de Lausanne*

Mandado a traducir al romance castellano en 1241 por Fernando III, el Fuero Juzgo se conserva en treinta y dos manuscritos medievales y diez de época posterior. La mayoría de los manuscritos conservados –incluso los códices castellanos– presenta rasgos dialectales occidentales, hasta tal punto que cabe imaginar que existiera una versión leonesa anterior a la traducción fernandina. Aunque demostrar su existencia no es el objetivo principal de esta investigación, no puede ignorarse el fuerte componente occidental de un texto que, paradójicamente, la RAE editó en 1815 para promover la historia del castellano.

Con el apoyo de programas de alineamiento textual, se ha realizado la *collatio* de una ley del octavo libro (8.2.1) con el fin de estudiar las relaciones de parentesco existentes entre los manuscritos e intentar construir el *stemma codicum* del texto. Aunque se trata de un trabajo en curso y las conclusiones a las que hemos llegado son provisionales, se han podido identificar las diferentes familias de testimonios con relativa certeza. De hecho, unos sondeos en otras dos leyes (3.5.7 y 11.1.2) parecen confirmar algunas de las relaciones que la ley 8.2.1 nos indican.

El objetivo de esta presentación consiste en compartir los resultados de esta *collatio*, comentar algunos de los rasgos dialectales más sobresalientes de la tradición manuscrita del Fuero Juzgo y, sobre todo, ofrecer un panorama de lo que conocemos de esta tradición manuscrita tan compleja y de lo que queda por hacer para resolver las numerosas incógnitas que rodean la historia de este texto.

---

## El corpus LIMES: retos y problemas

Javier Rodríguez Molina  
*Universidad Complutense de Madrid*

El objetivo de esta comunicación es presentar el corpus LIMES, un corpus de textos del siglo XIV creado en el marco del proyecto LIMES, Límites y márgenes en la historia del español.

LIMES es un Corpus digital en proceso de elaboración que permite acceder a un conjunto de más de 50 testimonios medievales manuscritos redactados y copiados entre c1285 y c1425 mediante un triple sistema de edición: una transcripción paleográfica, una presentación crítica del texto y una imagen digital del códice. El corpus LIMES se ha elaborado con la plataforma web [TEI-TOK](#), una herramienta diseñada para desarrollar corpus en línea que combinan la marcación textual con la anotación lingüística (Janssen y Vaamonde 2020), y sus criterios de edición se basan en los de la red CHARTA. Una particularidad del corpus radica en el alto número de testimonios datados que acoge, 25 en total, hecho que convierte a este corpus en una herramienta de potencial interés para efectuar estudios lingüísticos sobre fuentes fiables (cf. Rodríguez Molina, Octavio de Toledo 2017).

En la presente comunicación se abordarán los problemas de datación, localización y selección de los manuscritos que nutren el corpus, se explorará el interés del corpus para el estudio de la lengua del siglo XIV, centuria escasamente abordada en la bibliografía previa y se mostrará la utilidad del corpus para avanzar en la datación de manuscritos medievales castellanos.

### Referencias bibliográficas

- Janssen, M.; Vaamonde, G. (2020). "Da edición dixital á análise lingüística. A creación de corpus históricos na plataforma TEITOK". Álvarez, R.; González Seoane, E. (eds.): *Calen barbas, falen cartas. A escrita en galego na Idade Moderna*. Santiago de Compostela: Consello da Cultura Galega. Colección Ensaio & Investigación, 271-292.
- Rodríguez Molina, J.; Octavio de Toledo, Á. (2017). "La imprescindible distinción ente texto y testimonio: el CORDE y los criterios de fiabilidad lingüística". *Scriptum Digital*, 6, 5-68.
- 

## El corpus *ineffabilis* y las lenguas y variedades iberorrománicas en el siglo XIX

Johannes Kabatek / Manuela Crivelli  
*Universität Zurich*

En 1854, al Papa Pío IX proclamó la bula “Ineffabilis Deus” acerca de la inmaculada concepción de la virgen María, un documento latino con argumentos teológicos y filosóficos cuya versión original latina tiene una extensión de 4476 palabras. En los años después de la proclamación, el sacerdote francés Marie-Dominique Sire encargó traducciones de la bula a más de 300 lenguas y variedades. El total de las traducciones, conservadas en la Biblioteca Vaticana, constituye un corpus paralelo de unas 1,8 millones de palabras.

El corpus contiene, además de una traducción al español y al portugués, una traducción al gallego, una al asturiano y varias al catalán: documentos de sumo interés para la reconstrucción de la historia de las lenguas iberorrománicas en la época de sus respectivos “renacimientos” y al mismo tiempo fuente interesante para la comparación interiberorrománica.

El objetivo de la presentación es doble: por un lado, se presentará una versión digitalizada de acceso abierto de las traducciones iberorrománicas de la bula, y por otro lado, mostraremos cómo las traducciones son ejemplos de elaboración lingüística: su traducción a lenguas sin tradición escrita a mediados del siglo XIX (o sin tradición escrita presente en la época por desconocimiento de los antecedentes medievales o, en todo caso, con escasa tradición escrita) supone un reto de creación de medios adecuados para ello. Aparte de cuestiones como la ortográfica lo que llama la atención es el grado de elaboración léxica y sintáctica, y veremos que aunque el texto base es el mismo, las diferentes versiones eligen en algunos casos estrategias bastante diferentes.

### Referencias bibliográficas

- Kabatek, J. (2003a). “¿En que consiste o *ausbau* dunha lingua?”. Álvarez de la Granja, M.a; González Seoane, E. (eds.): *A planificación do léxico galego*, Santiago de Compostela: Consello da Cultura Galega / Instituto da Lingua Galega, 37-51.
- Kabatek, J. (2003b): “‘*Estamos dando principio ahora á la gramática asturiana*’ – Louis Lucien Bonaparte, Manuel Fernández de Castro y la elaboración del asturiano escrito”. *Actas del I Conceyu Internacional de Lliteratura Asturiana*. Uviéu [Oviedo]: Academia de la Llingua Asturiana, 23-51.
- Kabatek, Johannes (2018): “Traducción y dignificación lingüística”. Del Rey Quesada, S.; Del Barrio de la Rosa, F.; González Gómez, J. (eds.), *Lenguas en contacto, ayer y hoy: traducción y variación desde una perspectiva filológica*. Frankfurt am Main: Peter Lang, 27-46.
- Sire, Marie-Dominique (1866): “Traduzione in tutte le lingue della Bolla Ineffabilis, colla quale Sua Santità il Papa Pio IX proclamò domma di fede l'Immacolata Concezione di Maria santissima”. *La Civiltà Cattolica*, 17/7/IV, 729-737.

---

Aula Berengo  
**Sesión Dialectología Histórica (2)**  
15:00 – 16:30

### Caracterización lingüística del corpus del español aguascalentense virreinal

Blanca Elena Sanz Martín  
*Universidad Autónoma de Aguascalientes*

Este trabajo analiza el español dentro del territorio que conforma el actual Estado de Aguascalientes (México) durante el periodo virreinal para determinar qué fenómenos lingüísticos se presentaron en ese lapso. Para ello, se documentaron textos cercanos a la oralidad, generados en ese territorio en los siglos XVI, XVII y XVIII y XIX (hasta 1821, cuando se consuma la Independencia de México).

Se toma como corpus los textos de los archivos históricos, pues para acercarnos a los registros orales de épocas anteriores es menester emplear textos que presenten la fluidez y espontaneidad asociada al habla cotidiana, como cartas privadas, denuncias, declaraciones, etcétera.

Otros trabajos del español en México virreinal son: la región del Altiplano Central (*Documentos Lingüísticos de la Nueva España. Altiplano-Central* de Concepción Company Company 1994), el Golfo de México (*Documentos Lingüísticos de la Nueva España. Golfo de México* de Chantal Melis y Agustín Rivero 2008) y Campeche (*Documentos Lingüísticos de la Nueva España. Provincia de Campeche* de Pedro Ángel Ramírez Quintana 2017).

## CONFORMACIÓN DEL CORPUS

Los textos analizados fueron obtenidos tanto del Archivo General de la Nación como del Archivo Histórico del Estado de Aguascalientes. El número de textos documentados por siglo fue:

Siglo XVI: 4 textos

Siglo XVII: 42 documentos

Siglo XVIII: 97 documentos

Siglo XIX: 12 documentos<sup>2</sup>

## RESULTADOS

Los fenómenos lingüísticos son de diversa naturaleza: fonológicos, morfológicos, sintácticos y léxicos.

Los fenómenos fonológicos observados son: seseo, betacismo, velarización, yeísmo, apertura y cierre de vocales, tendencia antihiática, sandhis, consonantes geminadas, cultismos del latín, sonorización o ensordecimiento de oclusivas, duplicación vocálica, aferésis o apócope, no vocalización de consonante en posición de coda, asimilación de <r> y <l>, metátesis y epéntesis.

En cuanto a los fenómenos morfológicos encontrados son los siguientes: límites entre palabras distintos a los de español actual, falta de empleo de contracciones como las del español contemporáneo (p. e. *del*) y paradigmas verbales distintos a la norma culta moderna.

En torno a los fenómenos sintácticos observados, éstos fueron los siguientes: duplicación de posesivos, determinante antes del nombre propio, leísmo, dequeísmo, uso de adverbios de cantidad como cuantificadores, uso de relativos y fórmulas de tratamiento específicas de la época.

---

<sup>2</sup> La escasez de documentos del siglo XVI se debe a que Aguascalientes se fundó hasta 1575 y durante el resto de ese siglo la población no superó los 100 habitantes. Para el siglo XIX, se toma una cantidad proporcional a los años que dura el Virreinato.

Respecto a los fenómenos léxicos, encontramos vocablos de origen náhuatl (p. e. *chichigua*).

## CONCLUSIONES

Este trabajo ofrece una primera caracterización del español en la región de Aguascalientes, por lo que su valor radica en que el análisis parte de textos representativos de la zona, con lo cual obtenemos valiosa información sobre el estado histórico de la lengua española en regiones de México que aún no han sido estudiadas de forma pormenorizada. Así, esta investigación contribuye a complementar los estudios diacrónicos y diatópicos del español mexicano.

---

### **La lengua, fiel reflejo de la sociedad: el caso de una transformación del léxico de la ciudad de Málaga en el siglo XIX a través de la prensa local**

Salvador Peláez Santamaría  
*Universidad de Málaga*

La presente propuesta de comunicación se encuentra enmarcada dentro del proyecto de investigación denominado *Lengua, Cultura y Sociedad en la Málaga del siglo XIX: documentos y ciudad*, del Ministerio de Economía y Competitividad (CODEMA-19, UMA-20, FEDERTA-123) del Gobierno de España y que tiene como finalidad formar un corpus integrado por documentos y textos malagueños del siglo XIX de muy diversa naturaleza, que nos permita formar un corpus ampliado mediante la marcación TEI (Text Encoding Initiative: <http://www.tei-c.org/index.xml>), indispensable para la integración con otros corpus electrónicos de naturaleza similar.

La ciudad de Málaga vive toda una transformación económica, política, tecnológica y, por tanto, social a lo largo del siglo XIX. Los preámbulos no fueron fáciles (independencia de las colonias, revolución francesa, guerra de independencia, pandemias...) pero gracias a la visión de futuro de una parte de la sociedad malagueña que vio en todas estas calamidades una oportunidad: aprovecharon el abaratamiento de las materias primas y apoyaron aquellos movimientos que tratan de impulsar el liberalismo en España, tras la muerte de Fernando VII. Se inicia una transformación tecnológica sin igual que dio lugar a la industrialización de la ciudad, la mejora en los transportes y de las comunicaciones, los servicios públicos, etc. que afectaron en mayor medida a las ciudades que crecieron en población y extensión (Fernández 2020). Todo ello provocó un cambio en la sociedad que puede observarse tras la lectura de la prensa local de la época: espectáculos, barcos y carruajes que parten o llegan a la ciudad desde el resto del país y del mundo; utensilios; enseres; enseñanza de idiomas; máquinas; medicamentos de nuevo uso, y un sinnúmero de novedades que llenan los comercios y bazares de la ciudad. Muchas de estas realidades no encuentran su reflejo en el habla local y, por tanto, la lengua usa sus recursos para significar esos nuevos conceptos a través de entradas nuevas, calcos semánticos u otros mecanismos lingüísticos (Torres 2021). Hacer una descripción de este nuevo corpus a través de la lectura de la prensa local malagueña (Fernández-Sanguino 2017) es el objetivo de la presente comunicación.



## Referencias bibliográficas

- Fernández Paradas, M. (coord.) (2020). "El gas". *Andalucía en la Historia. La ciudad moderna. Infraestructuras en red*. 68. Centro de Estudios Andaluces, 22-25.
- Fernández-Sanguino, J. (2017). *La prensa editada en Málaga durante la Guerra de la Independencia*. Servicio de publicaciones: Unicaja.
- Torres, M. (2021). "Sobre purismo lingüístico en la prensa española del siglo XIX: en torno a la figura de A. Capmany". *Anuario de estudios Filológicos*, 44, 227-296.

---

## De Cuervo al CORDIAM: los corpus lingüísticos en el contexto de la historia del español de América

Javier Medina López  
*Universidad de La Laguna*

El trabajo de Cuervo (1901) sobre su visión de conjunto del español americano es considerado como el comienzo de la indagación dialectológica hispanoamericana, aunque unos años antes el alemán R. Lenz ya había publicado algunos estudios referidos al español de Chile y la influencia araucana sobre el mismo. Con Cuervo, y otros contemporáneos, la investigación filológica se fundamenta en la época de los orígenes y en los contingentes humanos que llegaron al Nuevo Mundo. Sin embargo, importaban más algunas cuestiones referidas a la sincronía que aquellas que tenían que ver con la propia evolución diacrónica del español. El empleo de fuentes documentales literarias para América es el denominador común de todos los estudios de la época, salvo alguna alusión casi tangencial. Pero esta situación cambiará, precisamente, a partir de las posturas adoptadas por los investigadores al enfrentar teorías lingüísticas en un legendario debate oponiendo andalucismo frente a poligénesis. Esto provocará que se inicie un cambio de rumbo y se aporten tanto datos demográficos (cifras de pobladores), como indagaciones filológicas (publicación de documentos). De esta manera, se superará la contienda y se abrirán nuevas perspectivas de análisis basadas en la nueva realidad histórico-documental de la que dispondrán los investigadores. En este contexto, se debe en gran medida a la obra de Amado Alonso (1951) el ser el pionero en tratar la documentación histórica para su análisis del seseo y el ceceo. Sin embargo, aunque su artículo arrojó luz documental y ofreció a la polémica sobre la base constitutiva del español de América más de un punto de vista, los métodos inaugurados en la lingüística diacrónica no tuvieron una gran continuidad en su momento, dado que la mayor parte del interés de la Lingüística estaba centrada en la perspectiva estructural y sincrónica. No obstante, años después de publicar Alonso sus trabajos, una investigación pionera vino a abrir una puerta en el terreno del seseo. En efecto, el estudio de Cock Hincapié (1969), aunque novedoso en sus materiales, tuvo también una desigual continuidad. Habrá que esperar a la década de los años ochenta del siglo XX para ver cómo se publican corpus diacrónicos referidos al español americano, inicialmente con algunas limitaciones y, tiempo después, con una mayor dimensión hispánica (Castillo Lluch y Díez del Corral Areta 2019), situación que hoy en día puede decirse que cubre la mayor parte del continente de habla hispana. En este trabajo se presentará la contextualización de las primeras aportaciones documentales hasta llegar a la aparición del CORDIAM, plataforma que hoy en día supone un avance importantísimo en las fuentes para el estudio del español de América (Bertolotti y Company Company 2001). La importancia de estos corpus es clave para la comprensión global del devenir lingüístico del español en el Nuevo Mundo.

## Referencias bibliográficas

- Bertolotti, V.; Company Company, C. (2014). "El *Corpus diacrónico y diatópico del español de América (CORDIAM)*. Propuesta de tipología textual". *Cuadernos de la ALFAL*, 6, 130-148.
- Castillo Luch, M.; Diez del Corral Areta, E. (2019). *Reescribiendo la historia de la lengua española a partir de la edición de documentos*. Bern: Peter Lang.
- Cock Hincapié, O. (1969). *El seseo en el Nuevo Reino de Granada (1550-1650)*. Bogotá: Instituto Caro y Cuervo.
- Company Company, C. (2001). "Para una historia del español americano. La edición crítica de documentos coloniales de interés lingüístico". Funes, L.; Moure, J. L. (eds.), *Studia in honorem Germán Orduna*. Alcalá de Henares: Universidad de Alcalá, 207-224.
- Cuervo, R. J. (1901). "El castellano en América". *Bulletin Hispanique*, 3, 35-62.

---

Aula Baratto  
**Sesión Diseño de Corpus (4)**  
17:00 – 18:00

### **PHOR-in-One: Cross-linguistic PHonological and ORthographic similarity indices in four languages**

Ana Santos Costa / Ana Paula Soares / Montserrat Comesaña  
*Universidade do Minho / Universidad Nebrija*

The bulk of bilingual research has shown a processing advantage for translations that formally resemble one another (e.g., English-Portuguese "paper-pape"). To investigate the role of interlanguage form similarity in lexical processing, the degree of orthographic overlap of two translation equivalents is typically assessed using the normalized Levenshtein Distance (NLD), a metric that computes the number of insertions, deletions and substitutions required to transform one string into another, considering the length of both strings. Despite its relevance in the context of bilingual research, the difficulty in implementing a similar metric for phonology has largely discouraged the conduction of studies investigating the impact of phonological similarity in second-language acquisition and processing. To address this concern, here we introduce PHOR-in-One, a multilingual lexical database that offers a set of phonological and orthographic NLD estimates for 6,160 translation equivalents in European Portuguese, Spanish, German and American and British English, in a total of 30,800 words. Cross-language orthographic similarity indices were based on the classical NLD, whereas phonological overlap was estimated using a fine-tuned version of the NLD algorithm, where substitution costs vary as a function of the proximity of the source and target phonemes in the IPA feature space. PHOR-in-One also includes a new metric of phonographic NLD, an index of the degree of orthographic and phonological overlap combined, particularly useful for controlling for or manipulating both indices at once. In addition, it offers a comprehensive characterization of its lexical entries, including relevant information such as Part-of-Speech-dependent and independent frequency metrics, number of letters and phonemes, and phonetic transcription. PHOR-in-One is a comprehensive synchronic record of the relationship

between the sound patterns of four European languages, which combined with similar past and future inventories, may shed light on how these languages evolve together from a phonological, orthographic and phonographic perspective.

---

## Inteligencia artificial y variación en documentos históricos del proyecto CAFECOMMIEL

Ruth Miguel Franco / David Sánchez  
*Universitat de les Illes Balears*

El proyecto CAFECOMMIEL (Corpus Automático y Fenómenos de Contacto en Mallorca: Inteligencia, Entrenamiento y Lengua) tiene como objetivo la creación de un gran corpus multiformato del español en contacto con el catalán en Mallorca. CAFECOMMIEL se compone, en su etapa inicial, de siete corpus que abarcan desde epistolarios del siglo XVII hasta habla espontánea del XXI, lo que permitirá realizar estudios lingüísticos en varios ejes mediante la aplicación de herramientas informáticas.

El objetivo de esta comunicación es presentar los resultados de la aplicación de un algoritmo de agrupamiento a muestras de los documentos históricos del proyecto CAFECOMMIEL, en concreto, a textos del *Epistolario de Pedro de Santacilia* (s. XVII) y del *Corpus Mallorca* (s. XVIII). Estos documentos contienen información sobre su origen geográfico y su año de emisión, además de la tipología documental a la que pertenecen o el emisor o escribiente que los produjo, por lo que sus agrupaciones serán interpretables según varios parámetros. La metodología del estudio partirá de una ventana temporal concreta para llevar a cabo un análisis de componentes principales y un procedimiento jerárquico de agrupamiento. El método no solo arrojará el valor óptimo de regiones que describen la mayor parte de la variación, sino que permitirá buscar las causas de dicha variación: temática, léxica, gráfica o de cualquier otro tipo. En este trabajo se expondrán los hallazgos más relevantes y se caracterizarán los agrupamientos desde diferentes puntos de vista, con vistas a aportar datos sobre la creación de la variedad de español de Mallorca.

---

Aula Berengo  
**Sesión Sintaxis Histórica (4)**  
17:00 – 18:00

### Los demostrativos compuestos *aqueste* y *aquese*: nuevos datos para la cronología del cambio

María Heredia Mantis  
*Universidad de Huelva*

Los determinantes y pronombres demostrativos en su sistema medieval se dividían entre los simples (*este* y *ese*, y en sus orígenes el que llegara a ser artículo, *el*) y los compuestos reforzados con ECCE (*aqueste*, *aquese* y *aqueel*), que originalmente aportaban cierto valor pragmático de énfasis o “prominencia del referente” (Lapesa 2000: 354, Ranson 2005: 209), valor que ha sido puesto en duda recientemente por Enrique-Arias (2018: 1559). En el Siglo de Oro, señala Lapesa (1981: párr. 96.6) que los demostrativos *aqueste* y *aquese* seguían gozando de plena vitalidad junto con los

también demostrativos *estotro* y *essotro*. No obstante, los gramáticos Minshieu y Correas señalaban que a principios del XVII su uso era escaso. Así lo corroboró Girón Alconchel (2013: 864): en *La Celestina* ya se acusaba la pérdida, y en *El Quijote* aportaban valor arcaico al discurso de los personajes. Posteriormente, Ranson (2005:209) señalaba que los datos del *Corpus del español* apuntaban a la desaparición de *aquestos* en el siglo XVII.

Tras nuestros análisis de la obra del novelista sevillano Mateo Alemán (1547-1614) hallamos que el escritor utilizó profusamente estos demostrativos *aqueste* y *aquese*, y en menor medida también *estotro* y *essotro*. Estos datos son discordantes con las conclusiones del estudio dialectal de Enrique-Arias (2018:1563-66) en cuanto a *aqueste* frente a *este*, que demuestra que la desaparición de *aqueste* fue más rápida en Castilla que en Aragón, a través de los datos disponibles en CODEA+ entre los siglos XIII y XVI. A partir de este hallazgo, nos proponemos contrastar los resultados del estudio de la obra alemaniana con los datos que arrojan distintos corpus históricos para analizar la pervivencia de *aqueste* y *aquese* en los siglos XVI y XVII, y clarificar su proceso de desaparición. Entre los corpus que analizaremos están CORDE y CODEA+, que han sido fuente de estos estudios anteriores, y recurriremos a CORDIAM para evaluar si estos demostrativos llegaron a América.

### Referencias bibliográficas

- Enrique Arias, A. (2018). "Factores diatópicos en la variación entre "este" y "aqueste" en la historia del español". Arnal Purroy, M.<sup>a</sup> L. et al., *Actas del X Congreso Internacional de Historia de la Lengua Española: Zaragoza, 7-11 de septiembre de 2015*. Zaragoza: Institución Fernando el Católico, vol. 2, 1553-1569.
- Girón Alconchel, J. L. (2013): "Cambios gramaticales en los Siglos de Oro". Cano-Aguilar, R. (ed.), *Historia de la lengua española*. Barcelona: Ariel, 859-894.
- Lapesa, R. (1981). *Historia de la lengua española*. Madrid: Gredos.
- Lapesa, R. (2000). *Estudios de morfosintaxis histórica del español*. Cano-Aguilar, R.; Echenique Elizondo, M.<sup>a</sup> T. Madrid: Gredos, vol. 1.
- Ranson, Diana L. (2005): "Variation of the Spanish demonstratives *aqueste* and *este*". Wright, R.; Ricketts, P. (eds.), *Studies on Ibero-Romance linguistics dedicated to Ralph Penny*. Newark: Juan de la Cuesta, 187-214.

---

## The Asturian language and the evolution of Romance clausal structure

Avelino Corral Esteban  
*Universidad Autónoma de Madrid*

Over the past few years, the comprehensive study of syntactic variation across both the Germanic and the Romance branches in relation to the unmarked word order pattern has meant a more in-depth knowledge of the nature of the verb-second phenomenon – an extremely intricate typological concept because of the complex factors that give rise to such a word order restriction. The aim of this paper is to investigate word order phenomena in the Asturian language (Western Iberian: Spain) from a diachronic comparative perspective by examining the word order patterns found in a number of documents written over the course of two centuries, with a view to determining what constitutes the unmarked word order and gaining a better

understanding of its core syntax. Such a study has entailed collating research on the verb-first, verb-second and subject-verb-object orders.

The organization of this paper is as follows: after presenting and describing central concepts such as V1, V2 or SVO, providing a background to the Asturian language, and offering a succinct summary of the literary history of Asturian, including a presentation of the texts examined, and describing the methodological approach followed in this study, Section 5 offers an analysis of word order patterns in the aforementioned text material, Section 6 discusses the results and compares them with those obtained in the analysis of other Romance languages, and Section 7 provides a comprehensive analysis of the core clausal syntax and the pragmatics of the left periphery of Asturian with the hope of definitively determining whether it should be considered a V1, V2 or SVO language and discusses the plausibility of the results in relation to the diachronic development of Romance languages. Finally, the concluding section of the paper contains the study's main findings and a brief discussion on the evolution of Asturian and its historical relation to Spanish.

Among the most relevant findings of the study it is worth mentioning that the results of the study are compatible with both a V2 and a non-V2 grammar, which could perhaps be accounted for by the diachronic evolution of Romance languages in general. Strangely enough, Asturian appears to illustrate the three different typological labels, namely V1, V2, and SVO, simultaneously, since it still resembles vulgar postclassical Latin and early Old Romance in showing a high proportion of V1 patterns, which is mainly due to the placement of the informational focus in postverbal position. Additionally, its predominant V2 order appears to be a reflection of an intermediate stage during which it may have exhibited V2 property. Finally, the proportion of subject preverbal elements in V2 patterns appears to indicate that Asturian is gradually becoming closer to the SVO system, probably due to the influence of Standard Spanish. In conclusion, Asturian seems to represent a transitional point from the VSO syntax of vulgar Latin to the V2 syntax of medieval Romance languages.

## References

- Benincà, P. (2006). "A detailed map of the left periphery of Medieval Romance". Zanuttini, R.; Campos, H.; Herburger, E.; Portner, P. (eds.), *Crosslinguistic Research in Syntax and Semantics. Negation, Tense and Clausal Architecture*. Oxford: Oxford University Press, 53-86.
- Benincà, P.; Poletto, C. (2004). "Topic, focus, and V2". Rizzi, L. (ed.), *The Structure of CP and IP*. Oxford: Oxford University Press, 52-75.
- Biberauer, T. (2002). "Reconsidering embedded verb second. How "real" is this phenomenon?". *Working Papers in English and Applied Linguistics*, 8, 25-60.
- Bosson, G. (1984). "Diachronie und Pragmatik der Spanischen Wortstellung". *Zeitschrift für romanische Philologie*, 100, 92-111.
- Bosson, G. (1998). "La typologie des langues romanes". Holtus, G.; Metzeltin, M. Schmitt, C. (eds.), *Lexikon der Romanistischen Linguistik (LRL), vol. 7: Kontakt, Migration und Kunstsprachen. Kontrastivität, Klassifikation und Typologie*. Tübingen: Niemeyer, 1003-1019.
- Cinque, G. (ed.) (2002). *Functional Structure in DP and IP. The Cartography of Syntactic Structures, vol. 1*. Oxford: Oxford University Press.

- Frascarelli, M.; Hinterhölzl, R. (2007). "Types of topics in German and Italian". Winkler, S.; Schwabe, K. (eds.), *On information structure, meaning and form*. Amsterdam: John Benjamins, 87-116.
- Holmberg, A. (2012). "Verb second". Kiss, T.; Alexiadou, A. (eds), *Syntax – Theory and Analysis. An International Handbook, vol. 1*. Berlin: De Gruyter, 342-383.
- Jouitteau, M. (2010). "A typology of V2 with regard to V1 and second position phenomena. An introduction to the V1/V2 volume". *Lingua, Special issue: Verb-First, Verb- Second*, 120(2), 197-209 .
- Kaiser, G. A., (2002). *Verbstellung und Verbstellungswandel in den romanischen Sprachen*. Tübingen: Niemeyer.
- Ledgeway, A. (2017). "Late Latin verb second. The sentential word order of the *Itinerarium Egeriae*". *Catalan Journal of Linguistics*, 16, 163-216.
- Roberts, I. (1993). *Verbs and Diachronic Syntax. A Comparative History of English and French*. Dordrecht: Kluwer Academic Publishers.
- Sitaridou, I. (2012). "A comparative study of word order in Old Romance". *Folia Linguistica*, 46, 553-604.
- Wolfe, S. (2015). "The nature of Old Spanish verb second reconsidered". *Lingua*, 164, part A, 132-155.
- Wolfe, S. (2015). "Verb-initial orders in Old Romance: a comparative account". *Revue roumaine de linguistique*, 60 (2-3), 147-172.
-

Viernes, 7 de octubre de 2022

Aula Baratto

Sesión Sintaxis Histórica (5)

9:00 – 10:30

## La distribució dels clítics pronominals en els futurs i condicionals del català antic (segles XIII-XIV)

Aina Torres Latorre  
*Universitat de València*

En català antic, les formes de futur i condicional (FC), derivades de la perífrasi llatina CANTĀRE HABEŌ (Pérez-Saldanya 1998: 278), no s'havien gramaticalitzat completament. Convivien formes sintètiques, acompanyades per pronoms proclítics (*el faré*) o enclítics (*faré-lo*), i analítiques, amb mesoclisi (*fer-lo he*). Les formes analítiques no presenten univerbació (l'expressió gramaticalitzada no està aglutinada, cf. Lehmann 1985) i, així, estan menys gramaticalitzades. La posició dels pronoms respecte del verb depenia de factors sintàctics (Batllori *et al.* 2005; Francalanci *et al.* en premsa). L'enclisi i la mesoclisi comparteixen uns mateixos entorns i, per tant, comparant-ne les freqüències és possible estudiar el grau de gramaticalització del FC.

L'estudi parteix d'un corpus d'obres dels segles XIII i XIV extretes del CICA.<sup>3</sup> Les dades s'analitzen quantitativament segons entorns i es comparen amb resultats del castellà (cf. Bouzouita 2016b) i del navarroaragonès (cf. Primerano 2019), perquè sembla que la gramaticalització del FC s'avança a l'est de la península Ibèrica respecte l'oest (Bouzouita 2016a; Bouzouita & Sentí en premsa). S'espera que el català presenti un menor grau de mesoclisi que les altres dues llengües.

### Referències bibliogràfiques

- Batllori, M.; Iglésias, N.; Martins, A. M.<sup>a</sup> (2005), "Sintaxi dels clítics pronominals en català medieval". *Caplletra*, 38, 137-177.
- Bouzouita, M. (2016a). "La posposició pronominal con futuros y condicionales en el castellano medieval: ¿un caso de contacto de lenguas?". *El orden de palabras en las lenguas iberorrománicas medievales*, 29-juliol-2016, Universitat de Girona.
- Bouzouita, M. (2016b). "La posposició pronominal con futuros y condicionales en el código escurialense l.i.6: un examen de varias hipótesis morfosintácticas". Kabatek, J. (ed.), *Lingüística de corpus y lingüística iberorrománica*. Berlin: De Gruyter, 270-298.
- Bouzouita, M.; Sentí, A. (2022). "La gramaticalización del futuro y el condicional en el iberorromance del siglo XIV a partir de traducciones bíblicas paralelas: el caso del castellano y el catalán antiguos". Enrique-Arias, A. (ed.), *Traducción bíblica e historia de las lenguas iberorrománicas*. Berlin, De Gruyter, 243-274.

<sup>3</sup> Del XIII, *Vides de sants rosselloneses*, *Costums de Tortosa*, *Llibre de contemplació en Déu* (no extreta del CICA), *Clams i crims a la València medieval*, *Llibre de Cort de Justícia de Cocentaina*; del XIV, *Regiment de preservació de la pestilència*, *Crònica de Muntaner*, *Ordinacions con los reys e reynas d'Aragó*, *Epistolari de la València medieval*, *Un llibre reial mallorquí del segle XIV*, *Procés criminal contra Antònia Marquès*, *Manual de consells de Xàtiva*, *El procés de Sueca*.

CICA = Torruella, J. (dir.). Pérez-Saldanya, M.; Martines, J. *Corpus Informatitzat del Català Antic*. <http://www.cica.cat/>.

Francalanci, L.; Martins, A. M.<sup>a</sup>; Ordóñez, F.; Ribera, J. E. (en premsa). “Els pronoms febles (I): forma, posició i ordre”. Martines, J.; Pérez-Saldanya, M.; Rigau, G., *Gramàtica del català antic*.

Lehmann, C. (1985). “Grammaticalization: Synchronic variation and diachronic change”. *Lingua e stile*, 20, 303-318.

Pérez-Saldanya, M. (1998). *Del llatí al català. Morfosintaxi verbal històrica*. València: Publicacions de la Universitat de València.

Primerano, A. (2019). *La gramaticalización de los futuros y condicionales en el navarroaragonés de los siglos xiii y xiv: análisis morfosintáctico-pragmático*. TFM, Universitat de Gent.

---

## La distribución de los clíticos pronominales en catalán antiguo

Andreu Sentí / Miriam Bouzouita  
*Universitat de València / Humboldt-Universität zu Berlin*

Se ha propuesto que la posición de los clíticos pronominales del catalán antiguo está motivada por restricciones pragmático-sintácticas (Fischer 2002, Batllori *et al.* 2005). Los clíticos suelen situarse en proclisis en oraciones subordinadas (1a), mientras que otros entornos, como detrás de un sujeto, admiten variación (1b-c).

- (1) a. lo tort que **li teníem** (*Fets*:5)
- b. E ell **fià·s** tant en él (*Fets*:12)
- c. E Nostre Seyor **la amà** tant (*Fets*:11)

Además, los clíticos con verbos en futuro o condicional (FC) presentan otra particularidad: existe una forma analítica (mesoclisís-(2)), que refleja la gramaticalización incompleta de la perífrasis latina [infinitivo+HABERE].

- (2) E yo **mostrar-lo-us hé** a matar (*Fets*:72)

Para el castellano (ss. XIII-XIV), se ha demostrado que estas formas analíticas ocurren en los mismos contextos que las estructuras sintéticas con pronombres posverbiales (cf. Bouzouita 2011; Castillo Lluch 2002), que podrían interpretarse como formas con un mayor grado de gramaticalización (univerbación, cf. Lehmann 2020).

En esta comunicación presentaremos un análisis detallado de la distribución pragmático-sintáctica de los clíticos pronominales en el *Llibre dels Fets*, con el refuerzo del CICA como corpus de apoyo. Esto permitirá confirmar la hipótesis de que el comportamiento de los clíticos con FC es equivalente al del resto de los tiempos verbales y, por tanto, que la forma analítica del FC no es una estrategia de focalización en catalán (cf. Company 2006), sino una forma de posposición requerida por el contexto pragmático-sintáctico. Además, analizaremos la variación entre la mesoclisís y la enclisis para los FC, de acuerdo con los datos cuantitativos aportados por el corpus, puesto que no hay estudios previos que la expliquen. De hecho, estos datos cuantitativos nos permitirán concretar las diferencias diatópicas en el avance de la gramaticalización: la forma sintética de FC con enclisis aparece con mayor frecuencia en catalán que en otras lenguas iberorromances medievales (de acuerdo con los estudios de corpus realizados: Bouzouita 2016; Primerano & Bouzouita entregado). Finalmente, el análisis pretende recabar datos para explicar el proceso de cambio



lingüístico en la posición preverbal de los clíticos en los siglos XIII-XIV y analizar qué elementos relacionados con la enfatización lo favorecen. Para ello, este estudio incorpora el *Llibre dels Fets*, un texto historiográfico del siglo XIII, a los datos disponibles hasta ahora en publicaciones anteriores sobre el siglo XIV basadas en textos bíblicos (Bouzouita & Sentí en prensa).

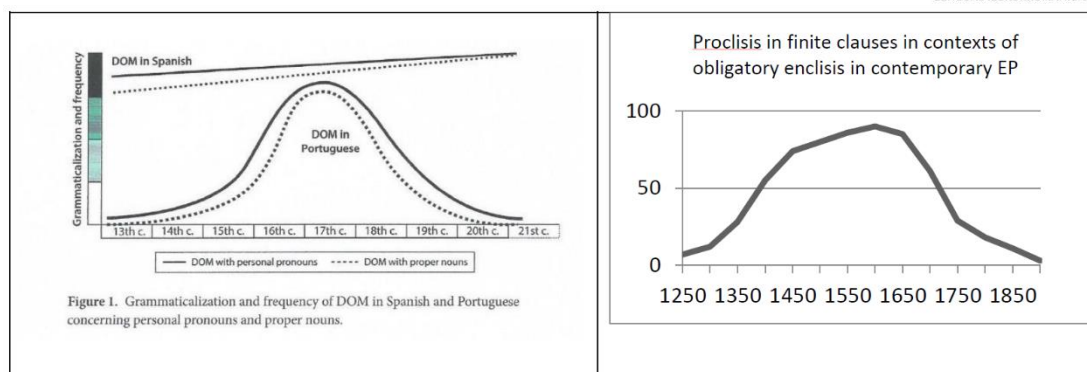
### Referencias bibliográficas

- Batllori, M.; Iglésias, N.; Martins, A. M.<sup>a</sup> (2005). "Sintaxi dels clítics pronominals en el català medieval". *Caplletra*, 38, 137-177.
- Bouzouita, M. (2011). "Future Constructions in Medieval Spanish: Mesoclisís Uncovered". Kempson, R. *et al.* (eds.), *The Dynamics of Lexical Interfaces*. CSLI, 89-130.
- Bouzouita, M. (2016). "La posposición pronominal con futuros y condicionales en el código escurialense I.i.6: un examen de varias hipótesis morfosintácticas". Kabatek, J. (ed.), *Lingüística de corpus y lingüística iberorrománica*. Berlin: De Gruyter, 270-298.
- Bouzouita, M.; Sentí, A. (2022) "La gramaticalización del futuro y el condicional en el iberorromance del siglo XIV a partir de traducciones bíblicas paralelas: el caso del castellano y el catalán antiguos", Enrique-Arias, A. (ed.), *Traducción bíblica e historia de las lenguas iberorrománicas*. Berlin, De Gruyter, 243-274.
- Castillo Llach, M. (2002): "Distribución de las formas sintéticas y analíticas de futuro y condicional en español medieval". Echenique Elizondo, M.T.; Sánchez Méndez, J. (edd.), *Actas del V Congreso Internacional de Historia de la Lengua Española*. Madrid: Gredos, 541-550.
- Company, C. (2006). "Tiempos de formación romance II: Los futuros y condicionales". Company, C. (ed.), *Sintaxis histórica de la lengua española*. Fondo de Cultura Económica/UNAM, 349-422.
- Fischer, S. (2002). *The Catalan Clitic System. A Diachronic Perspective on its Syntax and Phonology*. Berlin: de Gruyter.
- Lehmann, Christian (2020). "Univerbation". *Folia Linguistica Historica*, 41(1), 205-252.
- Primerano, A.; Bouzouita, M. (entregado): "La gramaticalización de los futuros y condicionales en el navarroaragonés de los siglos XIII y XIV".

---

**Dedicated corpora for specific research questions: are bell-shaped changes real?**

Ana Maria Martins  
*Universidade de Lisboa*



The above, left figure is taken from Döhla (2014), who sees the evolution of differential object marking (DOM) in Portuguese as consequence of the convergence towards Spanish in the 16<sup>th</sup>/17<sup>th</sup> century period (Portugal was under Spanish ruling between 1580 and 1640, but the strong cultural influence of Spain upon the Portuguese literate upper-class started much earlier and did not immediately disappear after 1640). There is an alternative interpretation for Döhla's (2014) figure, which is that there was never a bell-shaped change in Portuguese with respect to DOM, but instead two grammars in competition during the period in which the prestige/literate standard converged towards Spanish whereas nothing was happening in the grammar of common people relative to DOM and other phenomena (see Postma's 2010 on *failed changes*). In this talk I will focus mainly on another case of syntactic change in Portuguese that appears to display a similar bell-shaped curve (see the above, right figure), and I will propose that because extant texts are largely a product of the prestige grammar, they mask the complexity of the linguistic situation of the relevant historical period and create the illusion of a line of continuity between different objects. The subject to be addressed is clitic placement in finite clauses, specifically in those contexts where in 13<sup>th</sup> century as in contemporary European Portuguese enclisis is the regular position of clitic pronouns (this excludes negative and subordinate clauses, as well as affirmative clauses where quantifiers, *wh*- phrases and certain adverbs precede the verb). Proclisis, however, is the highly dominant pattern of clitic placement in the relevant contexts through the 16<sup>th</sup>-17<sup>th</sup> c. period. Textual evidence shows an increase in proclisis from the 13<sup>th</sup> to the 17<sup>th</sup> c. followed by a quasi-symmetric decrease. As for Spanish, it displays the same increase in proclisis as Portuguese until the 17<sup>th</sup> c. but stabilizes generalized proclisis in finite clauses from then on.

Most studies on the evolution of clitic placement in Portuguese have assumed a traditional perspective on linguistic change and provide accounts for syntactic change in one single grammar (e.g. Martins 1994; Galves/Brito/Sousa 2005, among others). Can we make a case for grammar competition between two social dialects instead? Not an easy task because current theoretical models can easily accommodate the bell-shaped change (as well as the wide variation in clitic placement among 16<sup>th</sup>/17<sup>th</sup> century writers and texts). Therefore, syntactic theory cannot by itself be invoked to exclude the single grammar approach in favor of the grammar competition approach. Based on the tradition of Romance Philology, I take here the view that certain types of texts are more likely to provide historical access to non-prestige grammars (recall, for example, the restricted set of generally accepted textual sources for the study of Vulgar Latin). I will therefore advocate the need for a dedicated corpus to address the relevant above-identified research question and will discuss data taken from a few selected textual sources, namely: texts of the 16<sup>th</sup> century playwright Gil Vicente, who can be thought of as 'the Portuguese Plautus' (concretely, the lines of 48 popular characters extracted from 9 of Vicente's plays); an anonymous 16<sup>th</sup> century kitchen book (Barros 2013), which is a collection of recipes assembled from different

(unidentified) sources but displays quite uniform 16<sup>th</sup> century orthographic and grammatical features; a 16<sup>th</sup>/17<sup>th</sup> century book of medicinal recipes (Barros 2017), which has similar features as the kitchen book. The data offered by these specific texts are significantly different from those found in most texts of the same period (see above, right figure), and support, in my view, the hypothesis of grammar competition. If the proposal I will present is on the right track, it will be a demonstration that large amounts of data may not be good enough. Instead, in particular cases, we may need a small-sized judiciously selected corpus. DOM will be also investigated in the same texts, in order to check whether it behaves as predicted by Döhla's figure above or differently.

## References

- Barros, A. Leal de, ed. (2013). *As Receitas de Cozinha de Um Frade Português do Século XVI*. Imprensa da Universidade de Coimbra.
- Barros, A. Leal de (ed.) (2017), *Remédios vários e receitas aprovadas; segredos vários. Edição do Caderno II do manuscrito 141 do Arquivo Distrital de Braga*. Imprensa da Universidade de Coimbra.
- CET (Centro de Estudos de Teatro), *Teatro de Autores Portugueses do Século XVI*. <http://cet-e-quinientos.com/autores>.
- Döhla, H.-J. (2014). "Diachronic convergence and divergence in differential object marking between Spanish and Portuguese", Braunmüller, K.; Höder, S.; Kühl, K. (eds.), *Stability and Divergence in Language Contact. Factors and Mechanisms*. Amsterdam/Philadelphia: John Benjamins, 265-289.
- Galves, C.; Britto, H.; Paixão de Sousa, M. C. (2005). "The change in clitic placement from Classical Portuguese to Modern European Portuguese: Results from the Tycho Brahe Corpus". *Journal of Portuguese Linguistics*, 4(1), 39-67.
- Martins, A. M. (1994). *Clíticos na História do Português*. PhD dissertation, Universidade de Lisboa.
- Postma, G. (2010). "The impact of failed changes". Breitbarth, A.; Lucas, C.; Watts, S.; Willis, D. (eds.), *Continuity and Change in Grammar*. Amsterdam/Philadelphia: John Benjamins, 269-302.

---

Aula Berengo  
**Sesión Diseño de Corpus (5)**  
9:00 – 10:30

## **Oralidad y escrituralidad en los corpus paralelos de las lenguas romances: el proyecto DiacOralEs**

Araceli López Serena / Santiago del Rey Quesada  
*Universidad de Sevilla*

En la lingüística histórica hispánica, el modelo concepcional de Koch y Oesterreicher (1990 [2007]) ha despertado gran interés sobre todo en lo que concierne al análisis de elementos lingüísticos susceptibles de ser considerados característicos de la lengua oral que se traslucen en textos del pasado (cf. Oesterreicher 2004). Dichos

elementos pueden ser presuntamente auténticos –como los que se asocian a los discursos escritos por semicultos (cf. Fernández Alcaide 2009) o los que se incluyen en actas inquisitoriales (cf. Eberenz 2003)–, o recreados en la literatura dialógica, teatral, novelesca, etc., de diversas épocas (cf. López Serena 2007; Del Rey 2019). Si bien la lingüística de corpus puede ayudar a reconocer algunos de estos elementos, la determinación de los fenómenos de la inmediatez comunicativa que puedan postularse como plausibles en textos del pasado no puede ser ajena de un escrutinio minucioso del cotexto y del contexto cultural y literario en el que se producen los discursos, lo que conduce, inevitablemente, a la necesidad de considerar rasgos propios de la distancia comunicativa en convivencia con rasgos propios de la distancia en los testimonios escritos.

El propósito de esta comunicación es doble. Por un lado, se describirán los presupuestos teórico-metodológicos que tutelan el proyecto DiacOralEs, al que pertenecen los autores de esta comunicación, sobre las relaciones entre oralidad y escrituralidad en la historia de las lenguas romances y se reflexionará sobre cómo la lingüística de corpus puede contribuir a conseguir los objetivos de este proyecto y de qué manera se pueden conciliar los intereses de la lingüística cuantitativa con los de la investigación filológica y textual. Por otro, se presentarán los dos corpus –uno de literatura dialógica romance (sss. XV-XVII) y otro de textos periodísticos (ss. XVII-XIX)–, a la vez paralelos y comparables, que pretendemos generar en el marco de DiacOralEs, incluidas las guías metodológicas en que se basará su confección –transcripción en XML, etiquetado en TEI, alineación de los textos paralelos, visualización y motor de búsqueda, etc.– con la finalidad de plantear las dudas metodológicas y los problemas de marcación que se nos plantean en la fase inicial del proyecto.

### Referencias bibliográficas

- Del Rey Quesada, S. (2019). “Variantes de la *oralidad elaborada* en la segunda mitad del siglo XIX: dos traducciones coetáneas de *Los cautivos* de Plauto”. *Oralia*, 22(2), 283-326.
- Eberenz, R.; de la Torre, M. (2003). *Conversaciones estrechamente vigiladas. Interacción coloquial y español oral en las actas inquisitoriales de los siglos xv a xvii*. Lausana/Zaragoza: Hispanica Helvetica/Pórtico.
- Fernández Alcaide, M. (2009). *Cartas de particulares en Indias del siglo xvi: edición y estudio discursivo*. Madrid/Frankfurt am Main: Iberoamericana/ Vervuert.
- Koch, P.; Oesterreicher W. (1990[2007]). *Lengua hablada en la Rumania: español, francés e italiano*. Versión española de Araceli López Serena. Madrid: Gredos.
- López Serena, A. (2007). *Oralidad y escrituralidad en la recreación literaria del español coloquial*. Madrid: Gredos.
- Oesterreicher, W. (2004). “Textos entre inmediatez y distancia comunicativas. El problema de lo hablado escrito en el Siglo de Oro”. Cano, R. (coord.), *Historia de la lengua española*. Barcelona: Ariel, 729-769.
-

## Normalización ortográfica y explotación de corpus: el caso de ODE

Miguel Calderón Campos / Gael Vaamonde  
*Universidad de Granada*

La recuperación de información a partir de un corpus lingüístico depende en gran medida de la anotación que se haya realizado previamente. Es obvio que un corpus lematizado y anotado morfosintácticamente podrá ser interrogado sobre más —y más abstractas— cuestiones que un corpus al que no se le haya añadido información gramatical alguna, lo que convierte la anotación del corpus en un factor determinante de su explotación posterior y, en definitiva, de su utilidad como herramienta de investigación (Rojo 2021: 113). Sin embargo, el número de corpus diacrónicos anotados es todavía escaso, debido a la alta variación ortográfica que suelen presentar los textos históricos y, especialmente, las transcripciones de documentación manuscrita. Esta particularidad reduce el porcentaje de acierto de los anotadores automáticos y obliga, en último término, a dedicar mayor tiempo y esfuerzo a la revisión manual. Es por ello que la normalización ortográfica de los textos como paso previo a la anotación lingüística se ha convertido en una solución cada vez más habitual en la construcción de corpus históricos. No en vano, en los últimos años se han desarrollado herramientas que facilitan esta tarea en mayor o menor medida (Baron & Rayson 2008, Feliciano de Faria et al. 2010).

En el corpus ODE (Oralia Diacrónica del Español) estamos construyendo un corpus histórico de la Edad Moderna mediante la herramienta TEITOK (Janssen 2014), que permite integrar múltiples niveles de edición del texto de forma secuencial (transcripción original, desarrollo de abreviaturas, normalización ortográfica, etiquetación morfosintáctica y lematización) y obtener información sobre ellos de forma relacional. Con esta estrategia, la normalización ortográfica del corpus no solo facilita la anotación de los niveles siguientes (i.e. las formas anotadas), sino que también permite establecer comparaciones con los niveles precedentes (i.e. las formas originales). En términos de explotación de corpus, esta última posibilidad conlleva dos ventajas adicionales que no suelen ofrecer los corpus históricos y que son objeto de análisis en la presente comunicación: (i) la obtención de variantes ortográficas de una misma forma normalizada, y (ii) la detección automáticamente de diferencias ortográficas con base fonológica, como son casos de seseo (*sinco* vs. *cinco*) y ceceo (*cauza* vs. *causa*), así como aspectos relacionados con el vocalismo átono (*sepoltura* vs. *sepultura*) o con las consonantes líquidas (*corchon* vs. *colchón*). Finalmente, ODE permite proyectar esta información en un mapa, constituyendo así un recurso especialmente útil para la investigación en dialectología histórica.

### Referencias bibliográficas

- Baron, A.; Rayson, P. (2008). "VARD 2: A tool for dealing with spelling variation in historical corpora". *Proceedings of the Postgraduate Conference in Corpus Linguistics*, Birmingham, Aston University.
- Feliciano de Faria, P.; Picasso, F.; Kepler, N.; Paixão de Sousa, M. C. (2010). "An integrated tool for annotating historical corpora". *Proceedings of the Fourth Linguistic Annotation Workshop*. Uppsala: Association for Computational Linguistics, 217-221.
- Janssen, M. (2014). *TEITOK, a Tokenized TEI environment*. <http://teitok.corpuswiki.org/site/index.php> (accessed 25 January 2022).

Rojo, G. (2021). *Introducción a la lingüística de corpus en español*. London/New York: Routledge.

---

## **Directrices para la construcción de un corpus diacrónico de la oralidad**

Salvador Pons Bordería  
*Universitat de València*

El estudio de la diacronía del siglo XX plantea, entre sus especificidades, la posibilidad de acceder a registros de distintas microdiacronías, posibilidad esta vedada al estudio histórico de otras épocas. La construcción de un corpus diacrónico oral es, por esta razón, una oportunidad de investigación que plantea desafíos metodológicos específicos cuyos abordaje y resolución resultan un proceso necesario y previo a la recolección de grabaciones.

La presente comunicación abordará, de forma programática, los objetivos, problemas y posibles soluciones a los problemas asociados al tratamiento diacrónico de lo oral, sobre la base de la experiencia en las distintas versiones del corpus Val.Es.Co.

### **Referencias bibliográficas**

- Hobsbawm, E. (1994). *The age of extremes. The short twentieth century 1914-1991*. London: Michael Joseph.
- Llopis Cardona, A.; Pons Bordería, S. (2020). "La gramaticalización de *macho* y *tío/a* como ciclo semántico-pragmático". *Círculo de Lingüística Aplicada a la Comunicación*, 82, 151-16.
- Llopis Cardona, A. (forth.). "La gramaticalización del *igual* americano como reformulador y concesivo: ¿caso de réplica de *lo mismo*?". *Onomázein*, 56.
- Pons Bordería, S. (2014). "El siglo XX como diacronía: intuición y comprobación en el caso de *o sea*". *Rilce*, 30(3), 985-1016.
- Pons Bordería, S.; Llopis Cardona, A. (2020). "Some reflections on semantic-pragmatic cycles". *Journal of Historical Pragmatics*, 21, 2.
- 

Aula Baratto  
**Sesión Sociolingüística Histórica (3)**  
11:00 – 12:30

## **Polifuncionalidad de *como* en un corpus del español rioplatense**

Laura D. Ferrari  
*Universidad Nacional General Sarmiento / Universidad de Buenos Aires*

Ante la pregunta planteada por Fleischman (1999) y retomada por otros autores de si los marcadores discursivos desarrollan un funcionamiento paralelo a las categorías gramaticales, esta autora presenta un estudio de caso sobre los marcadores *like* en inglés y *genre* en francés para contribuir al estudio de la pragmática comparativa y concluye que no existen diferencias cualitativas relevantes entre los desarrollos de

los marcadores discursivos y las categorías gramaticales. En esta misma línea Mihatsch (2010) analiza el cambio funcional de la partícula *como*, que transforma el elemento comparativo en una partícula de aproximación, sin dejar de analizar los cambios sintácticos que este proceso acarrea, como así también los valores semántico-pragmáticos que este aproximador adquiere en los variados usos en los corpus que analiza. La autora sostiene que en muchas lenguas existen nexos comparativos que se convierten en marcadores de aproximación, pero que para que eso ocurra, es necesario el reanálisis de la partícula comparativa.

En el ámbito del español peninsular, Fuentes Rodríguez y Gutiérrez Ordoñez (2019) en un avance hacia la formulación de una sintaxis del discurso proponen el término de macrosintaxis y la enmarcan dentro de la lingüística pragmática. Uno de los fundamentos de los estudios macrosintácticos es la existencia de categorías específicas que poseen un funcionamiento determinado en el nivel del enunciado, como por ejemplo, los operadores y los conectores.

En esta comunicación, mi objetivo es identificar y analizar en un corpus de español rioplatense el funcionamiento de *como* y su variante *como que*, en su evolución de nexos comparativos a marcadores discursivos de aproximación y caracterizar los valores pragmático-discursivos que presentan, así como su naturaleza categorial. El corpus sincrónico está constituido por emisiones elicidadas a partir de entrevistas semidirigidas. Los informantes son hablantes del español rioplatense, dos hombres y dos mujeres con instrucción superior completa o en curso y en un rango de edad que va desde los 24 a los 34 años.<sup>4</sup>

### Referencias bibliográficas

- Fleischman, S. (1999). "Pragmatic markers in comparative perspective: A contribution to cross-language pragmatics". Paper presented at PRAGMA 99. Tel Aviv, June 1999.
- Fuentes Rodríguez, C.; Gutiérrez Ordoñez, S. (eds.) (2019). *Avances en Macrosintaxis*. Madrid: Arco Libros.
- Mihatsch, W. (2010). "Sincronía y diacronía del aproximador *como*". *Revista Internacional de Lingüística Iberoamericana*, 8, 2(16), 175-201.

---

### Léxico del cuerpo humano en Colombia, análisis sincrónico y diacrónico en las variedades del español caqueteño y sucreño

Johnatan E. Bonilla  
*Universiteit Gent*

La presente ponencia expone los resultados de un piloto de recolección de datos léxicos del campo semántico Cuerpo Humano mediante técnicas de *crowdsourcing* para el Nuevo Atlas Lingüístico-Antropológico de Colombia (NALAC) en los departamentos de Sucre y Caquetá. Los datos recolectados se constituyen como una base de datos espacial-no relacional integrada al Sistema de Información Geográfica del Instituto Caro y Cuervo (SIGICC). El SIGICC es una herramienta para la administración, consulta y análisis de información lingüística geolocalizada. Dado que los resultados del Atlas Lingüístico-Etnográfico de Colombia (ALEC) (Instituto Caro y

---

<sup>4</sup> Las grabaciones fueron realizadas en el LIS (Laboratorio de Investigaciones Sensoriales), Conicet, Buenos Aires, Argentina.

Cuervo, 1981) fueron digitalizados y puestos en línea como atlas web<sup>5</sup> y base de datos (Bonilla et al. 2020) integrada al SIGICC es posible el mapeado de los resultados y su comparación con los datos recogidos entre 1956 y 1978. Ahora bien, el proyecto NALAC tiene dos propósitos, por una parte documentar y caracterizar el español de manera sincrónica en las regiones no visitadas por el ALEC (Orinoquía, Amazonía, Insular) y, por otra, permitir un análisis diacrónico del cambio lingüístico en el país en un rango de 50 a 70 años. En este sentido, los cuestionarios lingüísticos de ambos proyectos, ALEC y NALAC, guardan un grado de similitud que permite la comparación y se expande en fenómenos o entradas no antes documentadas. Sin embargo, en cuanto a la selección de hablantes se separa del ALEC, que basó su muestra en informantes tipo *NORMs*<sup>6</sup>, y opta por una visión pluridimensional teniendo en cuenta una distribución por rangos de edad y género. Para este piloto fueron seleccionados dos departamentos explorados durante los trabajos del ALEC. Sucre, de la región Caribe, documentado inicialmente en 1959 y con una encuesta suplementaria en 1977 (Flórez, 1983) y Caquetá, parte de la región Amazónica, del que solo se documentaron dos municipios en 1975 (*Ibid.*). Para la recolección de información se implementó un aplicativo web de recolección de texto y audio a partir de preguntas prediseñadas e imágenes y se trabajó de la mano con estudiantes y profesores de la Universidad de la Amazonía en Florencia, Caquetá y de la Universidad de Sucre en Sincelejo, Sucre cuyas edades oscilan entre 18 y 61 años. En general, la presentación resumirá los principios teórico-metodológicos de la investigación NALAC, la metodología de recolección de datos utilizada, sus principales beneficios y debilidades y, por último, el análisis, tanto sincrónico como diacrónico, de la variación léxica recolectada sobre el cuerpo humano en las variedades del español caqueteño y sucrense de Colombia.

### Referencias bibliográficas

- Bonilla, J.; López, R.; Llanos, A.; Bejarano, D.; Bernal, J. A. (2020). "Proyecto de digitalización y nuevas perspectivas tecnológicas del Atlas Lingüístico-Etnográfico de Colombia". Gallego, A.; Roca Urgel, F. (coords), *Dialectología Digital Del Español*, col. *Verba: Anuario Galego de Filoloxía*. n. 80, Universidade de Santiago de Compostela, Servicio de Publicaciones, 13-28.
- Chambers, J. K.; Trudgill, P. (1998). *Dialectology*. Cambridge University Press.
- Flórez, L. (1983). *Manual del Atlas Lingüístico-Etnográfico de Colombia*. Bogotá: Instituto Caro y Cuervo.
- Instituto Caro y Cuervo (1981). *Atlas Lingüístico-Etnográfico de Colombia*, ALEC (Vols. 1-6). Bogotá: Instituto Caro y Cuervo.

---

### Diacronía del mensaje publicitario impreso en el s. XIX: de los primeros avisos de comienzos de siglo a la explosión de los anuncios comerciales del cierre de la centuria

Sara Robles Ávila  
*Universidad de Málaga*

---

<sup>5</sup> <https://alec.caroycuervo.gov.co>.

<sup>6</sup> Non-mobile old rural males (Chambers and Trudgill, 1980).



Basándonos en un corpus de más de dos millones de piezas impresas –que se va ampliando progresivamente– en diarios de la época objeto de estudio, el s. XIX, nuestra propuesta tratará de analizar la evolución a nivel lingüístico, pragmático y textual de aquellos primeros avisos de principios de siglo con los que los particulares y pequeños negocios ofrecían sus productos o servicios a la población, textos escritos en un lenguaje poco elaborado y sometidos a un patrón discursivo muy particular; pasando por estadios intermedios inestables, donde se aprecian claras huellas del pasado más inmediato y avances expresivos que inauguran una nueva fase; hasta llegar a los mensajes del último tercio del XIX, donde se constata un tratamiento intencional de los recursos verbales, más en sintonía con el lenguaje publicitarios prototípico que todos conocemos.

En nuestro recorrido diacrónico nos detendremos en el estudio de las tipologías textuales de estos mensajes y en su evolución desde propuestas meramente informativas y denotativas, hasta otras persuasivas, sugerentes y connotativas. Igualmente, analizaremos el despliegue progresivo de recursos de realce aplicados al producto o servicio que se anuncia; y de fórmulas de apelación al receptor para que se interese por aquellos que se ofrece y lo adquiera. Con todo ello intentaremos mostrar la imagen dinámica de un género discursivo que nace, crece y madura en un siglo fundamental para la empresa publicitaria en España: el s. XIX.

### Referencias bibliográficas

- Buzek, I.; Šinková, M. (eds.) (2015). *Una cercana diacronía opaca: estudios sobre el español del siglo XIX. Études romanes de Brno*, 36.
- Eguizábal Maza, R. (1998). *Historia de la publicidad*. Madrid: Editorial Eresmas / Celeste Ediciones.
- López Serena, A.; del Rey Quesada, S.; Carmona Yanes, E. (eds.) (2020). *Tradiciones discursivas y tradiciones idiomáticas en la historia del español moderno*. Berlin: Peter Lang.
- Montero, M.; Rodríguez, N.; Verdura, F. (2010). *De la nada al consumo*. Sevilla: Comunicación Social.
- Ramírez Luengo, J. L. (coord.) (2012). *Por sendas ignoradas. Estudios sobre la lengua española en el siglo XIX*. Lugo: Axac.
- Robles Ávila, S. (2004). *Realce y apelación en el lenguaje de la publicidad*. Madrid: Arco/Libros.
- Robles Ávila, S. (edit.) (2005). *Aspectos y Perspectivas del Lenguaje de la Publicidad*. Analecta Malacitana, Anejo LVI.
- Robles Ávila, S. (en prensa). "Aproximación al lenguaje publicitario en prensa en el s. XIX", Granada, Comares.
- Romero Gualda, M.<sup>a</sup> V. (2005). *Lenguaje publicitario*. Barcelona, Ariel.

Aula Berengo  
**Sesión Sintaxis Histórica (6)**  
11:00 – 12:30

**La influencia del lexema verbal en la gramaticalización del futuro y condicional iberorrománico: el navarroaragonés en los siglos XIII-XIV**

Antoine Primerano

*Universiteit Gent / Humboldt-Universität zu Berlin*

Los futuros y condicionales (FC) iberorrománicos medievales exhiben, en entornos sintáctico-pragmáticos idénticos, una variación entre formas analíticas con pronombre átono mesoclítico (p.ej. *cantar-lo-é*) y sintéticas con pronombre pospuesto (p.ej. *cantaré-lo*) (cf. Bouzouita 2016, *inter alia*). Las segundas estructuras parecen reflejar un avance de la gramaticalización, puesto que se observa la univerbación (cf. Lehmann 2020) entre los dos componentes de la perífrasis latina original ([infinitivo + HABERE]). Los estudios previos dedicados a esta variación han propuesto factores que parecen favorecer el uso de la posposición, como la presencia de síncope verbal (p.ej. *avré* vs. *averé*), el condicional (frente al futuro) y la presencia de un verbo no finito detrás del FC (p.ej. *podré lo fazer* vs. *diré lo Ø*) (cf. Matute, Pato 2010; Bouzouita 2016; Primerano, Bouzouita entr.).

El estudio de corpus realizado en la presente investigación revela, además, que la posposición parece ocurrir sobre todo con ciertos verbos (*decir, haber, hacer, ser*, entre otros), por lo cual el lexema también parece perfilarse como una variable importante en la variación. Así, esta comunicación se enfocará en este patrón e intentará explicar por qué estos verbos son particularmente propensos a admitir pronombres pospuestos. Se han extraído de la parte navarroaragonesa (ss. XIII-XIV) de la base de datos OSTA (Gago, Pueyo 2021) todas las formas de FC de 62 verbos diferentes, documentados previamente con pronombre mesoclítico y/o pospuesto por Primerano & Bouzouita (entr.). Para cada verbo se ha calculado (i) su frecuencia de *tokens*, en cualquier tiempo y modo verbal, en el corpus entero, y (ii) la frecuencia de determinadas configuraciones en sus formas de FC (sincopado vs. asincopado, futuro vs. condicional, con verbo no finito o no). A partir de estos datos constatamos que la tendencia a la posposición en ciertos verbos resulta de (i) una alta frecuencia de *tokens* de estos lexemas, y (ii) su empleo frecuente en forma sincopada, condicional y seguidos por un verbo no finito, independientemente de la presencia de un pronombre átono. Estos resultados corroboran las afirmaciones de Bybee (2010) de que la alta frecuencia de los ítems y de los entornos favorecedores del cambio puede acelerar el ritmo y la expansión del mismo.

### Referencias bibliográficas

- Bouzouita, M. (2016). "La posposición pronominal con futuros y condicionales en el código escurialense I.i.6: un examen de varias hipótesis morfosintácticas". Kabatek, J. (ed.), *Lingüística de corpus y lingüística histórica iberorrománica*. Berlín: W. de Gruyter, 271-301.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: CUP.
- Gago Jover, F.; Pueyo Mena, F. J. (2021). *Old Spanish Textual Archive*. Hispanic Seminary of Medieval Studies. <http://osta.oldspanishtextualarchive.org>.

- Matute, C.; Pato, E. (2010). "Morfología y sintaxis en el código Escorial I.i.6". Enrique-Arias, A. (ed.), *La Biblia Escorial I.i.6. Transcripción y estudios*. Logroño: Biblias Hispánicas, 45-66.
- Lehmann, C. (2020). "Univerbation". *Folia Linguistica Historica*, 41, 205-252.
- Primerano, A.; Bouzouita, M. (entregado). "La gramaticalización de los futuros y condicionales en el navarroaragonés de los siglos XIII y XIV: una comparación con el castellano medieval". *Revue Romane*.

---

## Intensificación y énfasis en castellano medieval: la traducción de los infinitivos absolutos de la Biblia Hebrea

Andrés Enrique-Arias / Katharina Gerhalter

*Universitat de les Illes Balears / Karl-Franzens-Universität Graz*

El uso del infinitivo absoluto (o infinitivo tautológico) seguido por una forma flexionada del mismo verbo es una construcción característica del hebreo bíblico, que ha sido ampliamente estudiada (entre otros, Harbour 1999; Kim 2009; Hatav 2017, 2020). Se trata de una construcción opcional, con funciones pragmáticas como la aserción enfática, la intensificación o el refuerzo expresivo del significado verbal (Kim 2009). Mientras que los infinitivos absolutos tautológicos se reemplazan por otras estructuras en la *Vulgata* (sustantivo, participio, adjetivo, adverbio o expresión adverbial, reduplicación del verbo conjugado, paráfrasis) o simplemente se omiten (Kim 2009), en algunas traducciones castellanas medievales (hechas por traductores judíos), la estructura se calca literalmente: por ejemplo, *morir morrá* (Pérez Alonso 1997: 18; Viejo Sánchez 2006; Schwarzwald 2010: 133).

Para analizar este fenómeno de manera exhaustiva, analizamos todos los ejemplos pertinentes encontrados en el corpus paralelo *Biblia Medieval*; en total, 474 pasajes en 432 versículos (los números coinciden *grosso modo* con los mencionados por autores como Kim 2009). Las traducciones de estos pasajes – incluyendo las de la *Vulgata* – fueron anotadas manualmente con 16 categorías distintas (infinitivo, gerundio, repetición del verbo, omisión, etc.).

Nuestro estudio hace aportaciones novedosas en dos vertientes: por un lado, contribuye a caracterizar la lengua de la traducción bíblica, prestando especial atención al grado de fidelidad al original y a los esperables fenómenos de interferencia del texto subyacente. Y por otro, aprovecha las ventajas de un corpus de traducciones para observar las diversas estrategias empleadas en el castellano medieval para transmitir los valores pragmáticos del infinitivo absoluto hebreo.

### Referencias bibliográficas

- Enrique-Arias, A. (ed.) (2008). *Biblia Medieval*. <http://www.bibliamedieval.es>.
- Enrique-Arias, A. (ed.) (2022). *Traducción bíblica e historia de las lenguas iberorrománicas*. Berlin/Boston: De Gruyter.
- Galmés de Fuentes, A. (1996). *Influencias sintácticas y estilísticas del árabe en la prosa medieval castellana*. 2. ed., corregida y aumentada. Madrid: Gredos.
- Hatav, G. (2017). "The infinitive absolute and topicalization of events in Biblical Hebrew". Mosak Moshavi, A.; Tania Notarius, T. (Hg.), *Advances in biblical*

- Hebrew linguistics. Data, methods, and analyses.* Winona Lake, Indiana: Eisenbrauns, 207-229.
- Hein, J. (2020). *Verb Doubling and Dummy Verb. Gap Avoidance Strategies in Verbal Fronting.* Berlin: De Gruyter.
- Kim, Y.-K. (2009). *The Function of the Tautological Infinitive in Classical Biblical Hebrew.* Boston: Brill.
- Pérez Alonso, M. I. (1997). "Contribución al estudio de los hebraísmos léxicos, semánticos y morfosintácticos en la Biblia Medieval Romanceada". *MEAH Sección hebrea*, 46, 11-22.
- Schwarzwald, O. (2010). "On the Jewish Nature of Medieval Spanish Biblical Translations Linguistic Differences between Medieval and Post-Exilic Spanish Translations of the Bible". *Sefarad*, 70(1), 117-140.
- Viejo Sánchez, M. L. (2006). "El influjo de las construcciones hebreas de infinitivo absoluto en las construcciones sintácticas del tipo *hablar hablará, veer vimos* en las biblias medievales romancesas". Bustos Tovar, J. Jesús de (ed.), *Actas del VI Congreso Internacional de Historia de la Lengua Española.* Madrid, 29 de septiembre-3 de octubre de 2003. Madrid: Arco/Libros, vol. 2, 1203-1214.

---

## Un análisis de la pasiva en el español (pre)clásico basado en CODEA+2015

Bert Cornillie / Giulia Mazzola  
*KU Leuven*

El español moderno tiene dos construcciones pasivas: una pasiva perifrástica con el verbo copulativo *ser*, en (1), y una pasiva pronominal o refleja con *se*, ejemplificada en (2).

(1) Tampoco deben de ser tomadas estas frases, estos comentarios con mal fin, pues a veces no **son dichos** con mala idea (La Razón, 01/12/2004: Apoyo a la Guardia Civil, CREA).

(2) Ya era hora de que alguien se atreviera a decir algo sobre las barbaridades que **se dicen** de la Guardia Civil (examen escrito, CREA).

Si bien la pasiva perifrástica era la construcción más común hasta el siglo XV, hoy en día se restringe a contextos y estilos específicos y a menudo se considera como arcaica (cf. RAE, ASALE 2009: 784). En esta comunicación analizamos la actualización de la pasiva refleja a través de los documentos históricos de CODEA+2015 (GITHE 2015). En base a la idea de que el cambio lingüístico está mediado por las tradiciones textuales a varios niveles (Kabatek 2018, Rosemeyer en prensa), examinamos el papel de factores lingüísticos y socio-estilísticos en el desarrollo de las construcciones pasivas, para comprobar los efectos de las micro tradiciones discursivas (entornos semántico-pragmáticos) y las macro tradiciones discursivas (ámbitos textuales).

Los datos ( $n = 658$  ocurrencias) del corpus CODEA+2015 abarcan un periodo de cuatro siglos (de 1400 a 1799). A fin de examinar varios entornos semántico-pragmáticos, analizamos las formas pasivas en las oraciones subordinadas completivas dependientes de verbos matrices que expresan distintos actos de habla, tomando en consideración también la animación del sujeto (cf. Elvira 2002; Cabañas Maya 2005). Desde el punto de vista estilístico, consideramos los ámbitos discursivos del CODEA+2015 que representan el registro, e incluyen la comunicación

diplomática, eclesiástica, jurídica, municipal y privada (de más a menos formal). Además se toma en consideración el poder social del destinatario en comparación con el emisor (Brown, Gilman 1960), que a veces puede ser un predictor de distribuciones estilísticas particulares de variantes gramaticales.

Por lo que concierne el nivel semántico-pragmático, nuestro análisis cuantitativo demuestra que los actos de habla directivos (petición y mandato) son tierra propicia para la pasiva refleja. Asimismo, las frecuencias indican que prototípicamente el paciente de la pasiva refleja es un sujeto con un bajo nivel de animacidad. Contrariamente, la pasiva perifrástica se especializa en contextos en los que el paciente es humano o con animacidad alta.

Desde el punto de vista socioestilístico, por un lado, nuestros datos indican que contrariamente a lo esperado, el público de los documentos no afecta la alternancia de la pasiva. Por el otro lado, la actualización de la pasiva refleja arranca en los documentos formales de la diplomacia para ampliarse pronto a otros tipos de documentos. Así, a finales del siglo XVIII, la pasiva refleja ya había perdido su rasgo de formalidad y se convirtió en la construcción por defecto para expresar la voz pasiva. Paralelamente, la pasiva perifrástica se hacía cada vez más infrecuente e improductiva, limitándose a documentos conservadores como los textos jurídicos.

### Referencias bibliográficas

- Brown, R.; Gilman, A. (1960). "The Pronouns of Power and Solidarity". Sebeok, T. (ed.), *Style in Language*. Cambridge/Mass: MIT Press, 253-276.
- Cabañas Maya, R. (2005). *Análisis diacrónico de la construcción pasiva en el español*. Tesis doctoral. México: UNAM.
- Elvira, J. (2002). "Sobre el desarrollo de la pasiva refleja en español medieval". Echenique Elizondo, M.<sup>a</sup> T.; Sánchez Méndez, J. P.; Satorre Grau, F. J. (eds.), *Actas del V Congreso Internacional de Historia de la Lengua Española*, Valencia, 31 de enero-4 de febrero de 2000. Madrid: Gredos, 597-608.
- GITHE (2015). Codea+2015. *Corpus de documentos españoles anteriores a 1800*. <http://www.corpuscodea.es/> [consultado 10/02/2022].
- Kabatek, J. (2018). *Lingüística coseriana, lingüística histórica, tradiciones discursivas*. Madrid/Frankfurt am Main: Vervuert Verlagsgesellschaft.
- Real Academia Española. *Corpus de referencia del español actual*. <http://www.rae.es> [consultado: 22/02/2022]
- Real Academia Española; Asociación de Academias de la Lengua Española (eds.). (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Rosemeyer, M. (en prensa). "Discourse Traditions and Corpus Linguistics". Winter-Froemel, E.; Octavio de Toledo y Huerta, Á. (eds.), *Manual of Discourse Traditions in Romance* (Manuals of Romance Linguistics). Berlin: De Gruyter.

Aula Baratto  
**Sesión Dialectología Histórica (3)**  
15:00 – 16:30

**Los atlas lingüísticos como corpus para la investigación diacrónica de la lengua: calas para el estudio fonético, léxico y gramatical a partir de CORPAT**

Carolina Julià Luna / Carmen Quijada Van den Berghe / Assumpció Rost Bagudanch  
*UNED / Universidad de Salamanca / Universitat de les Illes Balears*

La aportación de los atlas lingüísticos al estudio histórico-etimológico de las lenguas es indiscutible (Rost 2013; Camus 2017; Gutiérrez 2021). La digitalización de los tradicionales mapas geolingüísticos y su publicación en línea —junto con bases de datos que permiten el acceso a los materiales que atesoran desde diferentes puntos de vista— supone un complemento valiosísimo a los corpus textuales y lexicográficos que permiten el desarrollo actual de la investigación sobre la evolución de las lenguas. En el ámbito iberorromance, las versiones digitales del *Atlas Lingüístico de la Península Ibérica*, el *Atlas Lingüístico y etnográfico de Castilla-La Mancha*, el *Atlas Dialectal de Madrid*, la *Cartografía Lingüística de Extremadura*, el *Atlas Lingüístico del Domini Català* y el *Petit Atlas Lingüístico del Domini Català* constituyen por el momento las únicas obras geolingüísticas de carácter tradicional disponibles en línea para el estudio histórico y contrastivo de la distribución de los fenómenos de cambio lingüístico de las variedades románicas habladas en la península ibérica.

Con el fin de complementar los datos de estas herramientas y de tener una visión del conjunto de los datos recogidos por la geografía lingüística en el siglo XX, se ideó y diseñó el *Corpus de los Atlas Lingüísticos (CORPAT)*, que actualmente se halla en una fase inicial de su desarrollo e incluye los datos de nueve atlas lingüísticos de pequeño dominio (*ALEA*, *ALEANR*, *ALEICan*, *ALECant*, *ALCyL*, *ALeCMan*, *ADiM*, *ALBI*, *CaLiEx*). El cotejo de este corpus junto a la versión digital del *ALPI* «permitirá enmarcar científicamente [...] los contenidos de los atlas regionales que, contextualizados, cobrarán todo su valor» (García Mouton 2021: 330).

En la presente comunicación se mostrarán algunas de las potenciales aportaciones de esta herramienta al estudio histórico del español desde las perspectivas fonético-fonológica, léxico-semántica y gramatical. A partir del análisis de algunos de los datos que contiene el corpus sobre diferentes mapas, se ilustrará la distribución de determinados fenómenos lingüísticos en el espacio mediante mapas digitales; se contrastarán los datos con otras herramientas electrónicas existentes para el estudio histórico del español; y se examinarán los procesos lingüísticos que subyacen a la variación que recogen los mapas dialectales.

### Referencias bibliográficas

- Camus Bergareche, B. (2017). "Perfecto simple y perfecto compuesto en los atlas lingüísticos españoles". *Moenia*, 23, 285-316.
- García Mouton, P. (2021). "El *ALPI* y la utilidad de sus materiales". Arejita, A. (arg.): *Aniztasuna: hizkeren berba-mintzoak. Euskaltzaindiaren II. Nazioarteko Dialektologia Biltzarra / Diversidad: vocablos y voces de las variedades lingüísticas. II Congreso Internacional de Dialectología de la Real Academia de la Lengua Vasca*. Bilbao: Euskaltzaindia, Colección Iker, 41, 321-331.

Gutiérrez, C. (2021). "Notas etimológicas sobre *guizque* y sus variantes en iberorromance", *Revista de lexicografía*, 27, 47-64.

Rost, A. (2013). "La transcripción fonética en estudios dialectales: propuestas en el caso del yeísmo". *Revista de Filología Española*, 93(1), 165-192.

---

## Hacia un corpus diacrónico del español de Colombia basado en sus atlas lingüísticos sincrónicos

Daniel Eduardo Bejarano Bejarano  
*Università degli Studi di Verona*

La geografía lingüística se consolida con el *Atlas Linguistique de la France* (ALF) de Gilliéron (Montes Giraldo 1999: 540). Estos productos, los atlas lingüísticos, ofrecen una descripción del lenguaje fiel al habla popular y rural de los territorios en los que se concentran. Para ello rastrean, capturan y exponen el habla de informantes con características específicas según perspectivas tradicionales (como es el caso de NORMS (Chambers y Trudgill 1994: 84), y más novedosas con la inclusión de diversas variables (como el método pluridimensional (Cardona 2012)). Así, en un atlas es posible observar la forma que toma la lengua en el momento en el que las pesquisas para su construcción son realizadas, esto es, un estado sincrónico.

Los atlas lingüísticos *per se* son enormes fuentes de información aprovechable para el análisis y la descripción de la lengua en un lugar determinado, es decir, son aprovechables como recursos similares a los corpus lingüísticos (como sugieren Torruella 2021 o Bernal et al. 2018 para el ALEC). Esto se ha visto especialmente potenciado con la incursión cada vez mayor de la tecnología observable en la creación y uso de bases de datos a partir de los materiales de atlas lingüísticos, como los casos del ALPI (<http://www.alpi.csic.es/>) o el ALEC (<http://alec.caroycuervo.gov.co/alec/>) o la creación de sistemas para la captura o despliegue de datos del tipo *Dialectos del español* (<https://www.dialectosdelespanol.org/intro>) o el Sistema de Información Geográfica del ALEC (<http://atlasweb.caroycuervo.gov.co/>). Así, un atlas supera su proyección análoga inicial y se convierte en un material que va más allá de los mapas. Además, la superposición de objetos como estos construidos desde una perspectiva sincrónica en un mismo espacio puede aprovecharse para comparar distintos momentos de la lengua en búsqueda de pistas de cambio o variación.

Con base en lo anterior, esta comunicación discute y presenta la potencial construcción de un corpus con perspectiva diacrónica creado a base de los atlas lingüísticos sincrónicos del español de Colombia, esencialmente el *Atlas Lingüístico-Etnográfico de Colombia* (ALEC) (Instituto Caro y Cuervo, 1982) y el *Nuevo Atlas Lingüístico-Antropológico de Colombia por regiones* (NALAC) (actualmente en desarrollo). Se comenta sobre su potencial uso con ejemplos comparativos de datos de ambas fuentes y se discute sobre su consolidación a partir de la etapa clave de sistematización de los datos apoyada por el diseño y construcción de una base de datos aprovechable desde distintas perspectivas, como la lingüística de corpus y la geolingüística.

### Referencias bibliográficas

Bernal, J., Bonilla, J., Rubio, R., Bejarano, D.; Llanos, A. (2018). "Atlas Lingüístico – Etnográfico de Colombia Geolinguistic Corpus". *LILA '18, V. International Linguistic and Language Conference*. DAKAM, 128-141.

- Cardona, M. (2012). *Atlas lingüístico pluridimensional del español en Belice (ALEB) Nivel fonético*. Bergen: Bergen Language and Linguistics Studies.
- Chambers, J.; Trudgill, P. (1994). *La Dialectología*. Madrid: Visor Libros.
- Instituto Caro y Cuervo (1982). *Atlas Lingüístico-Etnográfico de Colombia*. Bogotá: Instituto Caro y Cuervo.
- Montes Giraldo, J. J. (1999). "La Dialectología". *Thesaurus: boletín del Instituto Caro y Cuervo*, 54(2), 533-542.
- Torruella, J. (2021). "Atlas, diccionarios y corpus: tres recursos lingüísticos en contraste". De Beni, M.; Hourani-Martín, D. (eds.). *Corpus y estudio diacrónico del discurso especializado en español*. Berlin: Peter Lang, 25-54.

---

## **Tempo e espazo: o uso dos atlas lingüísticos na investigación diacrónica do galego**

Xulio Sousa Fernández  
*Universidade de Santiago de Compostela*

Os atlas lingüísticos tradicionais, a pesar de seren documentacións insubstituíbles e moi valiosas da lingua oral, non acostuman considerarse fontes primarias nas investigacións diacrónicas das linguas. A estrutura, formato e escasa difusión destas obras conducen con frecuencia a que sexan obviadas como referencias de consulta polos investigadores.

Esta comunicación pretende mostrar o valor dos datos que albergan os atlas lingüísticos para o estudo do cambio lingüístico. Utilizaranse materiais de dous proxectos de xeolingüística realizados en dous momentos do século XX e que ofrecen datos sobre o dominio lingüístico galego, o *Atlas Lingüístico de la Península Ibérica* e o *Atlas Lingüístico Galego*. As variables estudadas pertencen ó ámbito da morfosintaxe (*nós/nosoutros, lle/lles, tu/ti*, etc.) e na análise móstrase o interese de considerar a distribución xeográfica das variantes na investigación da historia e uso das unidades lingüísticas suxeitas a variación e cambio. Os resultados son contratados con investigacións previas que utilizan os corpus históricos como fonte de datos para unha dialectoloxía histórica.

### **Referências Bibliográficas**

- Dalbera, J.-Ph. (2007). "Linguistic atlases: Objectives, methods, results, prospects". Kawaguchi, Y.; Takagaki, T.; Tomimori, N.; Tsuruga, Y. (eds.), *Corpus-Based Perspectives in Linguistics*. Amsterdam: John Benjamins, 39-54.
- Julià Luna, C. (2021). "Del atlas lingüístico tradicional al corpus geolingüístico digital". *Scriptum digital. Revista de corpus diacrònics i edició digital en Llengües iberoromàniques*, 10, 109-147.
- Kawaguchi, Y., Minegi, M.; Viereck, W. (2011). "Introduction". Kawaguchi, Y.; Minegishi, M.; Viereck, W. (eds.), *Corpus-based Analysis and Diachronic Linguistics*. Amsterdam: John Benjamins.
- Sánchez-Prieto Borja, P. (2018). "El corpus ALDICAM-CM. Geografía lingüística diacrónica de la Comunidad de Madrid". *CHIMERA*, 5(1), 69-75.



- Sousa, X.; Rodríguez, D. (2017). "Do espazo ó tempo: o cambio lingüístico en galego a partir da xeolingüística". Regueira Fernández, X. L.; Fernández Rei, E. (eds.), *Estudos sobre o cambio lingüístico no galego actual*. Santiago de Compostela: Consello da Cultura Galega, 241-269.
- Sousa, X. (2016). "A xeolingüística e o estudo da historia da linguas". Rodríguez Guerra, A. (ed.), *Lingüística histórica e dialectoloxía: coordenadas do cambio lingüístico*. Vigo: Servizo de Publicacións da Universidade de Vigo, 35-52.
- Viereck, W. (2011). "The Atlas Linguarum Europae: A Diachronic Analysis of its Data". Kawaguchi, Y.; Viereck, W.; Minegishi, M. (eds.), *Corpus-based analysis and diachronic linguistics*. Amsterdam: John Benjamins Amsterdam, 21-36.

Aula Berengo  
**Sesión Diseño de Corpus (6)**  
15:00 – 16:30

**El Corpus de Documentos Medievales de San Andrés de Espinareda.  
Presentación del corpus y principales fenómenos lingüísticos**

Vicente J. Marcet Rodríguez/ M.<sup>a</sup> Nieves Sánchez González de Herrero  
*Universidad de Salamanca*

Esta comunicación tiene por objeto principal la presentación del Corpus de Documentos Medievales de San Andrés de Espinareda (CODOMSAE) y de sus principales características lingüísticas. Se trata de un corpus integrado por más de quinientos documentos notariales referidos y conservados en el monasterio de San Andrés, en Vega de Espinareda (El Bierzo, León), y que en la actualidad se custodian en el Archivo Histórico Nacional, en la sección Clero Secular Regular, entre las carpetas 835 y 861. Los documentos están siendo transcritos y editados por miembros del grupo GEDHYTAS, según los criterios de transcripción para textos antiguos de la Red Internacional CHARTA. Para la configuración del corpus, se seleccionaron aquellos escritos total o mayoritariamente en romance, por lo que contiene documentos de los siglos XIII, XIV y XV, siendo los redactados en esta última centuria los más comunes. Temáticamente, en su mayor parte se trata de textos transaccionales que afectan al monasterio, especialmente cartas de compraventa, donaciones y aforamientos. En menor medida, encontramos también testamentos, traslados de cartas reales, certificaciones, nombramientos y sentencias.

Desde el punto de vista lingüístico, se trata de un corpus muy interesante, no solo para la observación de la evolución de los fenómenos lingüísticos y los usos escriturarios de las escribanías de la zona, sino también para el estudio del contacto entre lenguas, puesto que los documentos de San Andrés de Espinareda, por su situación geográfica en el centro de la comarca de El Bierzo, en la frontera tradicional entre el gallego y el asturiano-leonés, presenta rasgos propios de ambos romances. Igualmente, resulta muy interesante para observar la expansión del castellano y su propagación por los dominios del leonés occidental a lo largo de la Edad Media, y especialmente durante el siglo XV, cuando tienden a volverse menos frecuentes los rasgos característicamente leoneses, a la vez que se vuelven más habituales los fenómenos propios del castellano.

Entre los fenómenos que pretendemos analizar someramente se encuentran la diptongación, la presencia de diptongos decrecientes, la evolución de -LJ- y grupos análogos, la evolución de PL-, KL-, FL- y el tratamiento de los pronombres personales y los posesivos, entre otros.

---

## **Onomástica femenina según el Corpus de Documentos Medievales de San Andrés de Espinareda**

Manuel Nevot Navarro  
*Universidad de Salamanca*

La presente comunicación tiene por objeto el estudio de la antroponimia femenina registrada en el Corpus de Documentos Medievales de San Andrés de Espinareda (CODOMSAE). Para este análisis, se ha recurrido a los documentos procedentes del monasterio de San Andrés de Espinareda, ubicado en la localidad berciana de Vega de Espinareda (León), material conservado actualmente en el Archivo Historia Nacional, sección Clero Secular Regular. La selección ha sido transcrita por los miembros del Grupo de Investigación Reconocido GEDHYTAS, conforme a las normas CHARTA.

La información contenida en las cartas de fuero, donaciones o compraventa analizadas resulta de especial relevancia, por la constante presencia femenina 'singularizada' (con nombre propio) en estas fuentes, a pesar de parecer a menudo vinculada al hombre con los acostumbrados "mujer de", "hija de". Por la localización de este centro religioso, en las confluencias de los dominios lingüísticos del gallego, leonés y castellano, resulta destacable tanto la elección de unos nombres de pila, que cambian al albur de las épocas (según costumbre tradicional) o la lengua prevaleciente de la que se toma la onomástica. Por lo que respecta a los apellidos, se constata la pobreza de las cuñas familiares. Asimismo, al presentarse estos apellidos de forma abreviada, es digno de mención cómo los paleógrafos proceden a la reconstrucción de los patronímicos.

En resumen, con esta documentación medieval, se dan a conocer los nombres y los apellidos de las mujeres en el área leonesa del Bierzo, zona plurilingüe gallega, asturleonés y castellana, sin que estos rasgos sean determinantes en la caracterización de la onomástica de la mujer.

### **Referencias bibliográficas**

- Faure Sabater, R. (2008). *Diccionario de nombres propios*. Madrid: Espasa Calpe.
- Marcet Rodríguez, V. J. (2018). "Los nombres propios de persona en documentación tardomedieval castellana". Arnal Prroy, M.<sup>a</sup> L.; Castañer Martín, R. M.<sup>a</sup>; Enguita Utrilla, J. M.<sup>a</sup>; Lagüéns Gracia, V.; Martín Zorraquino, M.<sup>a</sup> A. (coords.), *Actas del X Congreso Internacional de Historia de la Lengua Española: Zaragoza, 7-11 de septiembre de 2015*. Zaragoza: Institución Fernando el Católico, vol. 2, 1197-1212.
- Pichel Gotérrez, R. (2010). "Contribución á braquigrafía do galego medieval. Procedementos abreviativos e transcendencia escriptolingüística". Cano López, P.; Cortiñas Ansoar, S.; Dieste Quiroga, B.; Fernández López, I.; Zas Varela, L. (eds.), *XXX Simposio Internacional de la Sociedad Española de Lingüística*

(SEL): 1-14 de febrero de 2010. Universidade de Santiago de Compostela: Unidixital.

---

## **El *Old Spanish Textual Archive*: posibilidades de explotación**

Francisco Gago-Jover  
*College of the Holy Cross*

En el año 2015 el Hispanic Seminary of Medieval Studies (HSMS) comenzó a elaborar el *Old Spanish Textual Archive* (OSTA), retomando la idea propuesta en 1978 por John Nitti para “the creation of the Old Spanish Archive (OSA).”

En la cuarta edición de CODILI (Alcalá de Henares, 2016), hicimos una primera descripción del proyecto del *Old Spanish Textual Archive*, centrando nuestra atención en las tres etapas de desarrollo del corpus: 1) arquitectura y contenidos del corpus, 2) lematización y etiquetado morfológico, 3) desarrollo de la interfaz de consulta.

A finales del año 2017 dimos por concluida la primera etapa y a partir de este momento comenzamos a trabajar de forma paralela en el etiquetado morfológico y lematización del corpus y en el desarrollo de la interfaz de consulta.

Tras evaluar diferentes programas—NeoTag (Janssen 2012), Stanford Log-linear Part-Of-Speech Tagger (Toutanova 2003) y FreeLing (Carreras *et al.* 2004)—nos decantamos por FreeLing, una herramienta de procesamiento del lenguaje natural que, además de permitir la relativamente fácil personalización y mejora de los módulos de análisis morfológico, afijación y análisis multi-palabra, ya contaba también con un módulo de etiquetado de textos medievales (Sánchez-Marco *et al.* 2011).

El desarrollo de la interfaz de consulta estuvo condicionado no solo por la arquitectura del corpus sino también por su integración con las herramientas de código abierto IMS Open Corpus Workbench (CWB), cuyo componente central, el procesador de consultas CQP (Corpus Query Processor), permite realizar consultas utilizando la coincidencia de patrones con expresiones regulares. Al mismo tiempo preparamos un extenso manual de consulta.

En junio de 2021, después de seis años de trabajo, anunciamos de forma oficial el lanzamiento del *Old Spanish Textual Archive* [<http://osta.oldspanishtextualarchive.org/>]. En esta comunicación nuestro objetivo es doble: mostrar a los investigadores las posibilidades que ofrece OSTA y delinear las siguientes etapas de trabajo: mejora del lematizado y etiquetado morfológico e incorporación de nuevos textos.

### **Referencias bibliográficas**

- Carreras, X. et al. (2004). “FreeLing: An Open-Source Suite of Language Analyzers”. *Proceedings of LREC'04*.
- Gago Jover, F.; Pueyo Mena, F. P. (2018). “El *Old Spanish Textual Archive*, diseño y desarrollo de un corpus de textos medievales: lematización y etiquetado gramatical”. *Scriptum Digital*, 7, 25-35.

- Gago Jover, F.; Pueyo Mena, F. P. (2018). "El *Old Spanish Textual Archive*, diseño y desarrollo de un corpus de textos medievales: el corpus textual". *Cuadernos del Instituto Historia de la Lengua*, 11, 165-209.
- Hardie, A. (2012). "CQPweb: Combining Power, Flexibility and Usability in a Corpus Analysis Tool." *International Journal of Corpus Linguistics* 17(3), 380-409.
- Janssen, M. (2012). "NeoTag: A POS Tagger for Grammatical Neologism Detection." *Proceedings of LREC'12*.
- Nitti, J. (1978). "Computers and the Old Spanish Dictionary". *Computers and the Humanities*, 12, 43-52.
- Sánchez-Marco, C. et al. (2011). "Extending the Tool, or How to Annotate Historical Language Varieties". *Proceedings of the 5th ACL-HLT*.
- Toutanova, K., et al. 2003. "Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network." *Proceedings of HLT-NAACL 2003*, 252-259.

---

Aula Baratto  
17:00 – 17:30

### **El *Atlas Histórico del Español* y la dialectología del español antiguo**

Andrés Enrique-Arias / Marina Gomila Albal / Florencio del Barrio  
*Universitat de les Illes Balears / Università Ca' Foscari-Venezia*

Desde que Ramón Menéndez Pidal asentara las bases científicas de la lingüística histórica hispánica hace más de un siglo, el estudio de la variación y el cambio lingüísticos en la historia del español ha incorporado interpretaciones de la evolución del castellano en relación con los dialectos románicos vecinos (Fernández-Ordóñez 2011). Pero, a pesar de los avances de todo tipo que se han dado a lo largo del último siglo, a día de hoy falta un estudio sistemático de conjunto sobre la variación lingüística peninsular. Resulta esencial delimitar claramente la extensión de los rasgos comunes y diferenciales mediante un estudio objetivo de la extensión geográfica de las estructuras lingüísticas para así poder entender la configuración dialectal del castellano respecto de los romances vecinos. Además, un planteamiento unitario de un gran número de fenómenos puede revelar interrelaciones entre estructuras de gran interés para la teoría gramatical.

En esta comunicación se presenta el proyecto del Atlas Histórico del Español (AHE), una herramienta de acceso libre en internet con la que se pretende avanzar en el conocimiento de la variación dialectal a lo largo de la historia del español. El Atlas propuesto consiste en una base de datos de rasgos estructurales (con atención preferente a la morfosintaxis en esta primera fase, pero que en su desarrollo incluirá igualmente los ámbitos gráfico-fónico, léxico y discursivo) formada a partir de la compilación de materiales de referencia y de la recolección de datos de corpus dialectales. Para cada una de las estructuras estudiadas los resultados se presentan en forma de un breve estado de la cuestión ilustrado con gráficos y mapas. A partir de los resultados del Atlas, el proyecto pretende contribuir al estudio de la variación y el cambio lingüístico desde una perspectiva dialectal, abordar fenómenos no estudiados desde esta perspectiva, y sobre todo, extraer generalizaciones significativas sobre el origen y la difusión geográfica de los cambios lingüísticos que

puedan extrapolarse a otras situaciones de contacto de lenguas y variedades. Con ello se pretende, además de profundizar en el estudio de estos fenómenos variables, avanzar en la consecución de una de las tareas pendientes de la lingüística histórica hispánica, la de constituir el mapa variacional del español a lo largo de sus diferentes etapas evolutivas, es decir, establecer una tipología de cuáles son los principales fenómenos sujetos a variación en cada época y área geográfica en la evolución del español (Del Barrio de la Rosa 2018).

### **Referencias bibliográficas**

Del Barrio de la Rosa, F. (2018) *Espacio variacional y cambio lingüístico en español*. Madrid: Visor Libros.

Fernández-Ordóñez, I. (2011) *La lengua de Castilla y la formación del español*. Madrid: Real Academia Española.

---