

Algunas posibilidades de investigación del  
proyecto *Procesos de gramaticalización en la historia  
del español* (Documentos PROGRAMES)

Daniel M. Sáez Rivera

*Universidad Complutense de Madrid*

*Proyecto PROGRAMES*

<https://www.ucm.es/procesosdegramaticalizacionenlahistoriadelespanol>

# Presentación e historia PROGRAMES

- **PROGRAMES = Procesos de gramaticalización en la historia del español** (grupo de investigación UCM y proyecto de investigación Ministerio de Educación y Ciencia > MinEco)
- Desde **2002, 5 ediciones.**
- Edición actual (2016-2019): *Programes 5: “Procesos de gramaticalización en la historia del español (V): **gramaticalización, lexicalización y análisis del discurso desde una perspectiva histórica**”* (MinEco: Ref. FFI2015-64080-P)
- **Investigador principal** del proyecto:
  1. José Luis Girón Alconchel (2002-2016)
  2. Francisco Javier Herrero Ruiz de Loizaga (2016-)
- **Página web:** creación en 2010 como <http://portal.ucm.es/web/programes> por Patricia Fernández y Daniel Sáez, migración en 2014 a <https://www.ucm.es/procesosdegramaticalizacionenlahistoriadelespanol>
- **Webmaster y técnico informático del proyecto:** Daniel M. Sáez Rivera

# Prehistoria PROGRAMES

- **Primeras transcripciones informáticas:** Word, transcripción diplomática (cuasifacsimilar) curso 1998-1999 por becarios de colaboración del Dpto. de Filología Española I bajo la dirección de José Luis Girón Alconchel:
  1. **Gutiérrez de los Ríos y Córdoba, Francisco, Conde de Fernán Núñez:** *El hombre practico o discursos varios sobre su conocimiento y enseñanzas*. Bruselas: Felipe Foppen, 1686 (por Álvaro S. Octavio de Toledo): <http://bdh-rd.bne.es/viewer.vm?id=0000115088&page=1>
  2. **Fernández, Marcos:** *Olla podrida a la española...* Amberes [=Amsterdam]: Felipe van Eyck, 1655 (por Daniel M. Sáez Rivera) <http://bdh-rd.bne.es/viewer.vm?id=0000088961&page=1>
- > estudio de la ortografía con ayuda del programa de corpus WordSmith: Sáez Rivera (2005)

# Historia PROGRAMES

- Transcripción y edición de SOBRINO, Francisco, y Félix A. de ALVARADO (2002): *Diálogos nuevos (1708) de Francisco Sobrino y Diálogo decimoquinto (1718) de Félix Antonio de Alvarado* (ed. e intro. de Daniel M. Sáez Rivera) [en línea], <http://parnaseo.uv.es/Lemir/Textos/Sobrino/Index.htm>
- Curso 2002-2003: comienza la **aplicación del TEI (Text Encoding Initiative)**, inspirado en los corpus académicos de CREA y CORDE (cf. Pino/Santalla, 1996), para la **transcripción y etiquetado manual de formato y fenómenos lingüísticos de interés** (etiquetado automático descartado porque se decide realizar transcripción paleográfica) en forma de [Documentos PROGRAMES](#)
- **Transcripción y etiquetado por colaboradores** (alumnos de grado, máster y doctorado; jóvenes doctores)
- **Extensión actual del corpus:** más de **un millón de palabras** (c. 1.300.000 palabras)
- **Diciembre 2016:** [consulta en línea del corpus](#) a través de [LYNEAL](#) (por Hiroto Ueda y Antonio Moreno Sandoval):

# Características Documentos PROGRAMES

- “La colección de textos que aquí se presenta tiene por objeto **ensanchar el número de textos disponibles para estudio de la historia de la lengua española en una edición o transcripción fiables filológicamente**, respecto a cuatro ámbitos en concreto, con frecuencia confluentes:

1) **épocas poco estudiadas de la historia del español**, como la transición del español clásico al moderno (1640-1726)

2) **una atención marcada a textos no literarios** y, al mismo tiempo, especialmente valiosos para calibrar el grado de acercamiento de la lengua escrita a lo que podría ser la lengua hablada de las distintas épocas: colecciones epistolares, relatos históricos, noticiosos y anecdóticos, publicaciones periódicas, ensayos divulgativos, textos técnicos, etc.

3) **tradiciones discursivas en las que apenas se ha reparado hasta ahora**, como los manuales de cartas, los textos técnicos o los textos jurídicos del primer tercio del siglo XVIII

4) **autores que han despertado poca atención en la historia de nuestra lengua**, como Feliu de la Peña, Joseph Barzia y Zambrana, Bartholomé Guerrero Ludeña, Luis Salazar y Castro, o Francisco Mariano Nifo y Cagigal

” (cf. Documentos PROGRAMES) [subrayado añadido]

# Características Documentos PROGRAMES

“ Aparte de que ningún texto ha sido incluido en ningún corpus lingüístico del español en red, **los testimonios textuales elegidos dentro del proyecto carecen** normalmente **de edición moderna**, o esta no sigue criterios de transcripción válidos para un estudio diacrónico, por haberse realizado en tal edición modernización gráfica, de puntuación e incluso de lengua. En cambio, nuestros **criterios de transcripción** han sido de fidelidad absoluta al texto original (disponibles abajo).” (cf. **Documentos PROGRAMES**)

“Dados los objetivos de investigación del proyecto en cuanto a lingüística diacrónica (gramaticalización, morfología histórica, sintaxis histórica, semántica histórica), el equipo decidió formular una serie de criterios de transcripción de los textos acorde con tales objetivos: se efectuará básicamente una **transcripción paleográfica** (grafías, acentuación y puntuación); en cuanto a las **abreviaturas**, se **conservarán las referidas a formas y fórmulas de tratamiento**, mientras que se **desarrollarán entre corchetes el resto**. Se acordó también **no separar palabras por renglones, ni siquiera al final de página**, ya que el programa WordSmith no reconoce las palabras segmentadas y se trata de una información lingüística poco o nada relevante para nuestro estudio.” [**negritas añadidas**]

# Características Documentos PROGRAMES

“Estos textos hasta ahora de difícil acceso, pues era necesario desplazarse a las bibliotecas en que se conservan para poder consultarlos (su signatura se ofrece en la ficha de cada texto: ver infra), se ofrecen en **dos versiones**:

1) **Versión para analizar con un programa de análisis de corpus, como *Wordsmith***, que ha sido el empleado en el proyecto. Por ello, el texto se ofrece en un archivo con extensión .doc (para que funcione mejor en Wordsmith, es recomendable cambiar el formato a .txt) y marcado manualmente por los transcritores con una serie de etiquetas del tipo SGML que indican el formato y algunos fenómenos lingüísticos de interés para los investigadores del proyecto. El conjunto de etiquetas se puede descargar abajo, también en formato .doc, aunque el programa solo reconoce el formato .txt (el presente portal no permite descargar archivos en .txt, por eso el usuario debe hacer la conversión a .txt manualmente, guardando el archivo con tal extensión, para que *Wordsmith* pueda reconocer las etiquetas o *tags*). Con el deseo de facilitar el posible intercambio de información con otros proyectos similares, se ha seguido el TEI (*Text Encoding Initiative*) habitual en este tipo de documentos.

2) **Versión para leer:** en formato pdf, se ofrece el texto para una más cómoda lectura, despojado prácticamente de etiquetas, habiendo dejado solo las imprescindibles de formato, como las que delimitan cada folio o página, por ejemplo del tipo <p. 1> al comienzo de la página 1 y </p. 1> a su fin, o etiquetas para cursiva (<curs> </curs>), márgenes, columnas, etc., así como la que señala erratas (<sic> </sic>).” (cf. [Documentos PROGRAMES](#))

# Características Documentos PROGRAMES

- **Etiquetas de fenómenos lingüísticos**

<cdp>, </cdp> Complemento directo preposicional

<conda>, </conda> Condicional analítico

<dcl>, </dcl> Duplicación clítica

<ft>, </ft> Forma/fórmula de tratamiento

<futa>, </futa> Futuro analítico

<lasml>, </lasml> Laísmo

<lesml>, </lesml> Leísmo

<losml>, </losml> Loísmo

<marc>, </marc> Marcador del discurso [etiquetamos los siguientes marcadores: *con que/conque, a saber, a pesar de (que), no obstante, de hecho, últimamente*]



# Características Documentos PROGRAMES

- **Etiquetas de fenómenos lingüísticos solo empleadas en los textos transcritos 2003-2005**

<cd> </cd>

Complemento directo

<ci> </ci>

Complemento indirecto

<ppas>, </ppas>

Participio de pasado

<ra> </ra>

Forma verbal en *-ra*

<reg> </reg>

Régimen preposicional de los

verbos

<rel> </rel>

Relativo

<relart> </relart>

Artículo + relativo

<relcomp> </relcomp>

Relativo compuesto

# Ej. de Concord (Wordsmith): Garau (1991): "le"

Concord

File Edit View Compute Settings Windows Help

N	Concordance	Set	Tag	Word #	Sen	Sen	Para	Para	lead	lead	Sec	Sec	File	%
54	, i sogá de esparto al pescuezo- se le leió su sentencia con meritos, abjuró	r76>		19.312	54645%		057%				057%		Garau 1691.txt	67%
55	. Estando en forma de penitente: se le leió su sentencia con meritos, abjuró	r79>		19.976	56217%		059%				039%		Garau 1691.txt	69%
56	. Estando en abito de penitente: se le leió su sentencia con meritos, abjuró	r80>		20.205	56817%		070%				070%		Garau 1691.txt	70%
57	con maior piedad zelo, i constancia le á añadido cultos de la maior	ig6>		4.177	17019%		014%				014%		Garau 1691.txt	15%
58	7. capitulo 15. en que parece se <dcl> le </dcl> puede dar la enora buena a	dcl>		24.248	67445%		034%				034%		Garau 1691.txt	83%
59	curs> Senec. i mientras no se duele; ni le duele, ni se emmienda. Assi estos	urs>		5.920	20839%		021%				021%		Garau 1691.txt	21%
60	. Vele, i discurra aora a quien <dcl> le </dcl> toque, el remedio eficaz de	dcl>		25.233	70053%		037%				037%		Garau 1691.txt	86%
61	fuerres, i vivas razones, que <lesm> le </lesm> izo motivar su noble piedad,	sm>		14.392	42142%		050%				050%		Garau 1691.txt	50%
62	delante de los ojos, de quien <lesm> le </lesm> leiere, con mas fidelidad, i	sm>		507	2224%		0 2%				0 2%		Garau 1691.txt	2%
63	, i como el impio huie sin que <lesm> le </lesm> persiga nadie: <curs> fugit	sm>		5.417	19475%		019%				019%		Garau 1691.txt	19%
64	encomendaran a Dios su alma, <lesm> le </lesm> ize azer vna breve	sm>		12.337	37731%		043%				043%		Garau 1691.txt	43%
65	, que oi meretissimamente <lesm> le </lesm> gobierna como Lugartiniente	sm>		9.914	33122%		034%				034%		Garau 1691.txt	35%
66	, si su mucha vmildad no <lesm> le </lesm> negára a la luz. </p> <p>	sm>		12.010	37292%		042%				042%		Garau 1691.txt	42%
67	a el </cdp> no solo no <dcl> <lesm> le </lesm> </dcl> convencieran, pero ni	sm>		15.005	43845%		052%				052%		Garau 1691.txt	52%
68	lesm> </dcl> convencieran, pero ni aun le izieran duda, como varias vezes me	sm>		15.010	43831%		052%				052%		Garau 1691.txt	52%
69	vez, solo que para maior luzimiento, se le añadió a mano drecha al entrar, vn	<ft>		10.286	33857%		036%				036%		Garau 1691.txt	36%
70	de IESVS, que por muchos titulos se le deve el titulo, que en la Feé Triunfante	<ft>		1.105	3559%		0 4%				0 4%		Garau 1691.txt	4%
71	, con todas sus demas prendas, le merecieron por nuevos titulos la	[III]>		857	2990%		0 3%				0 3%		Garau 1691.txt	4%
72	que se llamava como el <dcl> <lesm> le </lesm> </dcl> mandó, o que dejára	sm>		25.646	71655%		039%				039%		Garau 1691.txt	88%
73	, i extraordinaria erudicion, <lesm> le </lesm> eligió el <ft> Santo Tribunal	sm>		814	2940%		0 3%				0 3%		Garau 1691.txt	3%
74	discrecion, i prudencia de quien <lesm> le </lesm> regenta, que es su Ministro,	sm>		18.741	53179%		055%				035%		Garau 1691.txt	65%
75	, de IESVS, a quantos <lesm> le </lesm> assistiamos, </cdp> sino	sm>		14.977	43758%		052%				052%		Garau 1691.txt	52%
76	Divina que prevista aquella, <lesm> le </lesm> destinó como a otro ludás	sm>		15.076	43932%		052%				052%		Garau 1691.txt	52%
77	de vengadoras olas, no <dcl> <lesm> le </lesm> </dcl> an de sepultar en el	sm>		5.843	20659%		020%				020%		Garau 1691.txt	21%
78	<cdp> a Cristo, </cdp> <lesm> le </lesm> vendian, teniendo en el,	sm>		5.222	19118%		018%				018%		Garau 1691.txt	19%
79	de la Ballena, que a su pesar <lesm> le </lesm> condujo al termino del	sm>		6.196	21545%		021%				021%		Garau 1691.txt	22%
80	su nombre glorioso, o que no <lesm> le </lesm> profanára covarde. Mas	sm>		25.657	71693%		039%				039%		Garau 1691.txt	88%
81	<cdp> a IESVS. </cdp> Ni <dcl> le </dcl> bastó al Valls la estoica	dcl>		16.350	47311%		057%				057%		Garau 1691.txt	57%
82	se resistia a la luz de Dios, i <lesm> le </lesm> precipitó a la fatalidad de su	sm>		14.845	43376%		051%				051%		Garau 1691.txt	51%
83	muerte, que le quiere dar quien <lesm> le </lesm> tiene forzado a morir. El	sm>		16.129	46831%		056%				056%		Garau 1691.txt	56%

concordance collocates plot patterns clusters filenames follow up source text notes

# Otro ejemplo práctico: Garau

- Comparación de dos ediciones de un mismo texto para cotejar variantes gráficas (1741: RAE, *Ortografía de la lengua castellana*) y lingüísticas:
- **1691: Garau, Francisco:** *La fee triunfante en quatro autos celebrados en Mallorca por el Santo Oficio de la Inquisicion en que an salido ochenta, i ocho reos, i de treinta, i siete relaiados solo uvo tres pertinaces.* [Mallorca]: viuda Guasp. *Descripción física:* [16], 120 [i.e. 119] p.; 4°. *Ejemplar:* Madrid, BN 2/61699. [Transcripción de Lorena Núñez Pinero, 2015]
- **1755: Garau, Francisco:** *La Fee triunfante en quatro autos celebrados en Mallorca por el Santo Oficio de la Inquisición....* Mallorca: Ignacio Frau. *Descripción física:* 24, 175 p.; 8° (14 cm). *Ejemplar:* Madrid, BN 2/62172 (DGMICRO/26989). [Transcripción de Lorena Núñez Pinero, 2015]

- Uso creativo de herramientas creadas con otros propósitos:  
comparación de documentos mediante Word para comprobar fácilmente los cambios.

Garau fundido.doc [Modo de compatibilidad] - Microsoft Word uso no comercial

Inicio Insertar Diseño de página Referencias Correspondencia Revisar Vista Complementos

ABC Referencia Sugerencia de información en pantalla para traducción  
Ortografía y gramática Sinónimos Traducir Definir idioma Contar palabras

Revisión

Eliminar Anterior Siguiente Comentarios  
Nuevo comentario

Control de cambios Globos

Marcas mostradas finales Mostrar marcas Panel de revisiones Seguimiento

Rechazar Anterior Siguiente Cambios

Comparar Mostrar documentos de origen Proteger documento

Comparar... Comparar dos versiones de un documento (estilo jurídico).  
Combinar... Combina revisiones de varios autores en un documento único.

... como ~~pag66bis~~ ~~pag66bis~~ ~~relapso~~ ~~permanez~~, no se ~~executó~~ ~~executó~~ en ~~el~~ ~~el~~ sino la comun de relajado al ~~braçobrazo~~ ~~seglar~~, muriendo como los otros ~~aà~~ ~~las~~ ~~buchtas~~ de ~~vaun~~ ~~garrote~~, ~~iy~~ ~~despues~~ ~~incinerizados~~, ~~iy~~ ~~confiscados~~ sus bienes, por ~~eregeherege~~, apostata, judaizante, relapso, convicto, ~~iy~~ ~~confesso~~. </p>  
<p> 2. Miguel Valls de Campos negociante de oficio, natural desta Ciudad residente en Ciutadella de Menorca, de edad de treinta, ~~iy~~ ~~nueve~~ años reconciliado, ~~iy~~ ~~preso~~ segunda vez, por crimen de judaismo relapso. ~~Salio~~ ~~Salio~~ desde la Inquisicion con las mismas insignias que el passado mordaza, corozza, ~~iy~~ ~~capotillo~~ <pag43> <pag44> de llamas, ~~iy~~ ~~Cruz~~ verde en las manos; pero estando ~~iaya~~ en la Iglesia antes de oír su sentencia abrió los ojos, para abjurar sus errores con clarissimas señales de bien convertido, ~~iy~~ ~~mucho~~ consuelo de los ~~catolicos~~: ~~leioselecatolicos~~, ~~leiosele~~ su sentencia con meritos, ~~i~~ ~~fué~~ ~~fué~~ relajado ~~aà~~ la Justicia, ~~iy~~ ~~brazo~~ seglar con confiscacion de bienes por ~~eregeherege~~, apostata, ~~judaizante~~ ~~judaizante~~, relapso, ~~iy~~ ~~confesso~~. </p> </pag66bis>  
<pag67bis> <p> 3. ~~sic~~ ~~Francisca~~ </sic> ~~Francisca~~ Cortès muger de Onofre Aguiló Aguiló, de Onofre, negociante de ~~ofieio~~, ~~sic~~ ~~ficio~~ </sic> natural, ~~i~~ ~~vezinay~~ ~~vecina~~ desta Ciudad, de edad de quarenta ~~iy~~ ~~ocho~~ años, reconciliada, ~~i~~ ~~y~~ presa segunda vez por ~~judaizante~~ ~~judaizante~~ relapsa ~~Salio~~. ~~Salio~~ (como todos los que se siguen en este Auto) con corozza, Cruz verde, ~~iy~~ ~~capotillo~~ de llamas: ~~leioseleleiosele~~ su sentencia con meritos, ~~iy~~ ~~fué~~ relajada ~~aà~~ la ~~Justicia~~, ~~i~~ ~~braço~~ ~~Justicia~~, ~~y~~ ~~brazo~~ seglar con confiscacion de bienes por ~~eregeherege~~, apostata, judaizante, relapsa, convicta, ~~iy~~ ~~confessa~~. </p>  
<p> 4. ~~Catalina~~ ~~Cathalina~~ Pomar, viuda de Rafael ~~Marti~~ ~~Marti~~ menor, ~~aliás~~ ~~aliàs~~ del ~~arpa~~ ~~Arpa~~ de oficio botiguero; natural, ~~i~~ ~~vezina~~ ~~desta~~ ~~y~~ ~~vecina~~ ~~de~~ ~~esta~~ Ciudad, de edad de ~~desta~~ ~~i~~ ~~vez~~ ~~un~~ ~~años~~ ~~reconciliada~~, ~~iy~~ ~~presa~~ ~~segunda~~ ~~vez~~ ~~por~~ ~~Judaizante~~ ~~Judaizante~~

Página: 26 de 57 Palabras: 32.016

140%

19:17 12/01/2017

# Conclusión

- **Pese a lo limitado del equipo y de los medios, resultados de gran calidad y difusión** (por ej. Girón Alconchel y Sáez Rivera, coords., 2014).
- **Muchos textos disponibles aún por estudiar** (fuente útil para trabajos de fin de grado, máster y tesis, aparte de artículos, capítulos, libros), más fácil ahora gracias a LYNEAL y otras herramientas de corpus como [AntConc](#) (gratis).
- Intención de ahondar en la **historia de los discursos especializados** y de los **manuales epistolares**.
- **Mejoras informáticas: perfeccionar la interfaz** con ayuda de Hiroto Ueda y Ana Stulic y **agilizar con herramientas automáticas la transcripción y etiquetado lingüístico**.

# Bibliografía

- Garau, Francisco (1999)[1691]: *Inquisición en Mallorca : La fe triunfante : Los grandes Autos de Fe celebrados en Mallorca en 1691* (prólogo de Dr. Llorrenç Pérez Martínez; versión, estudio preliminar y glosario, Dr. Lleonard Muntaner i Mariano), 2.ª ed. (1.ª ed. 1984). Mallorca: Miquel Font.
- Girón Alconchel, José Luis, y Daniel M. Sáez Rivera (eds.) (2014): *Procesos de gramaticalización en la historia del español*, Madrid/Frankfurt am Main, Iberoamericana/Vervuert.
- Pino, Marta / Santalla, M.ª Paula (1996): “Codificación morfosintáctica de corpus en lenguaje SGML”. *Revista de la Sociedad Española de Procesamiento del Lenguaje Natural*, 20: 101-117.
- Sáez Rivera, Daniel M. (2005): «La *Olla podrida a la española* (1655) o los presuntos entretenimientos de un maestro de lenguas del siglo XVII», en M.ª Carmen Cazorla Vivas et al. (coords.), *Estudios de historia de la lengua e historiografía lingüística: Actas del III Congreso Nacional de la AJIHLE* (Jaén, 27, 28 y 29 de marzo de 2003), Madrid, AJIHLE, pp.457-467.
- Sobrino, Francisco, y Félix A. de Alvarado (2002): *Diálogos nuevos (1708) de Francisco Sobrino y Diálogo decimoquinto (1718) de Félix Antonio de Alvarado* (ed. e intro. de Daniel M. Sáez Rivera) [en línea],  
<<http://parnaseo.uv.es/Lemir/Textos/Sobrino/Index.htm>>

Muchas gracias  
[dansaez@ucm.es](mailto:dansaez@ucm.es)

---

**Agradecimientos**  
José Luis Girón Alconchel  
Equipo PROGRAMES  
(dirs. José Luis Girón y  
Fr. J. Herrero)  
Hiroto Ueda  
Lorena Núñez