

ASinEs: Prolegómenos de un atlas de la variación sintáctica del español

ASinEs: Prolegomena to an atlas of syntactic variation in Spanish

Alba Cerrudo

Universitat Autònoma de Barcelona
alba.cerrudo@gmail.com

Anna Pineda

Universitat Autònoma de Barcelona
anna.pineda@uab.cat

Ángel J. Gallego

Universitat Autònoma de Barcelona
angel.gallego@uab.cat

Francesc Roca

Universitat de Girona
francesc.roca@udg.edu

Resumen

En este artículo se presenta el ASinEs¹, una aplicación con formato de atlas dedicada al estudio sincrónico de la variación sintáctica de los geolectos del español. Este proyecto es innovador, ya que no existe ningún atlas dedicado exclusivamente a investigar la variación geolectal de la sintaxis de esta lengua. La versatilidad del ASinEs permite también el estudio de geolectos de otros estadios del español, así como los de otras lenguas con las que está actualmente en contacto. Todo ello proporciona una potente herramienta para la investigación en el campo de la variación de las lenguas románicas y no románicas (vasco, inglés, lenguas amerindias, etc.).

El desarrollo de este proyecto cuenta con la colaboración del *Centre de Lingüística Teòrica* (Universitat Autònoma de Barcelona), el *Centro IKER* con sede en Bayona (Francia) y la *Real Academia Española*.

Palabras clave

atlas, español, geolectos, sintaxis, variación

Abstract

This paper introduces the ASinEs¹, an atlas-based application devoted to the study of the syntactic variation of Spanish geolects. This project is groundbreaking, as there is no other atlas exclusively devoted to study the geolectal variation of geolectal variants of Spanish. Although ASinEs was originally conceived to explore the current geolects of Spanish, its flexibility allows it to study both the geolects of previous stages and the geolects of other close-by languages. This provides us with a powerful tool to study variation of both Romance and non-Romance languages (Basque, English, Amerindian languages, etc.).

This project is being developed in collaboration with the *Centre de Lingüística Teòrica* (Universitat Autònoma de Barcelona), the *IKER Center* at Bayonne (France), and the *Real Academia Española*.

¹<http://www.asines.org>

Keywords

atlas, Spanish, geolects, syntax, variation

1 Introducción

Tradicionalmente, los estudios de dialectología (o geolingüística) y variación lingüística se han centrado en fenómenos que pertenecen a los niveles léxico, fonético o morfológico (cf. Chambers & Trudgill (1980); Chambers & Schilling-Estes (2013); Labov (1994, 2001); Labov et al. (2006); Petyt (1980), entre otros), utilizando a tal efecto técnicas cuantitativas (estadísticas), de reconstrucción (diacrónica), y comparativas (cf. Campbell (2001) y referencias allí citadas).

Muchos de esos trabajos tienen en cuenta factores sociales y geográficos para explicar el cambio / variación, y han dado lugar a avances importantes en nuestra comprensión de fenómenos socio-lingüísticos como la diglosia, los *continuum* geolectales o las áreas de transición (el llamado “*Sprachbund*”).

Otro de los resultados de esta línea de acción fue una caracterización adecuada de unidades como el “fonema”, el “morfema” o el “rasgo distintivo” (introducidos por el estructuralismo europeo), lo cual permitió y facilitó la investigación basada en el trabajo de campo y condujo a estudios tipológicos como los de Greenberg (1963).

En este apartado, revisamos brevemente los antecedentes en los estudios de geolingüística del español, así como los avances que ha habido en las teorías sintácticas y de variación en las últimas décadas, los cuales han permitido plantear un cambio de tendencia en la investigación geolectal.

1.1 Antecedentes en el estudio del español

Los estudios sobre la variación geolectal, en el caso de lenguas como el español, se han centrado en los mismos dominios mencionados al inicio: el léxico, la fonética y la morfología (cf. Alvar (1996b,a); Fernández-Ordóñez (2011); García Mouton (1994); Kany (1945), entre otros). El énfasis puesto en dichos fenómenos puede verse en los diferentes atlas lingüísticos (p.ej., ALPI, ALEANR, ALBI, ALEA, ALECAN, ALCyL), que suelen representar los siguientes tipos de variación:

- **Léxica:** reflejado en Figura 1a, donde aparecen recogidas las respuestas extraídas del ALPI a la pregunta sobre el nombre de “la cría de la cabra”.
- **Morfológica:** podemos ver muestras en Figura 1b, donde se observan las variantes del pronombre de segunda persona del singular.
- **Fonética:** tenemos ilustraciones en Figura 1c, donde se ofrecen las soluciones fonéticas del sustantivo *tejón*, que proviene del latín TAXU.

La misma situación puede observarse en los estudios diacrónicos, donde son mayoritarios los trabajos sobre el léxico y la morfofonología (con alguna excepción; cf. Company Company (2006, 2009, 2014); Lapesa (2000)). La siguiente cita, extraída de Sánchez Lobato (1994), recoge la idea —extendida— de que la sintaxis apenas presenta variación geolectal:

Los rasgos lingüísticos, pues, más característicos del español americano —frente a la subnorma castellana— se encuentran en esa nueva coine surgida de Andalucía. Lo esencialmente autóctono del español de América se encuentra en su aliento, en su voz, es decir, en la entonación, en el ritmo y en el léxico, no en la morfología. *En la sintaxis no hay diferencias notables.*

[Sánchez Lobato (1994, pg. 560),
énfasis nuestro]

La opinión de Sánchez Lobato no constituye, en absoluto, un caso aislado. Otros autores enfatizan el predominio del interés por el léxico y la morfo-fonología en los estudios geolectales del español:

La falta de atención a los problemas de sintaxis dialectal por parte de la dialectología tradicional, el estructuralis-

25. CRÍA DE LA CABRA



(a) Variación Léxica

27. PRONOMBRE 2ª P. PL.



(b) Variación Morfológica

15. Consonante en TAXU



(c) Variación Fonética

Figura 1: Ejemplos de fenómenos estudiados en atlas del español (casos tomados del ALPI y discutidos en Fernández-Ordóñez (2011, 2014))

mo y la gramática generativa está dejando un vacío en nuestro conocimiento de la realidad viva de las lenguas que tardará muchos años en llenarse. Son varias las razones de esta situación contra la que se viene clamando desde hace algunos años, sin que los intentos de ponerle remedio pasen de ser esfuerzos aislados. La primera causa me parece enteramente imputable a la dialectología que, desde sus orígenes, ha tendido a obviar este tipo de asuntos, posiblemente

porque los métodos que fue perfilando para la recolección de materiales fueran poco aptos para obtener información sobre la sintaxis de las hablas estudiadas. Por eso es totalmente acertada la observación de Gregorio Salvador de que «la sintaxis nunca ha sido hasta ahora ocupación seria de dialectólogos, sino de filólogos».

[Morillo-Velarde Pérez (1992, pgs. 219–220), énfasis nuestro]

Pueden leerse palabras similares en relación al dominio diacrónico:

Es casi un lugar común la afirmación de que los estudios de sintaxis histórica del español presentan un desarrollo muy limitado. Y aunque se ha publicado una bibliografía de sintaxis histórica con 548 títulos (Narbona 1984–1985), lo cierto es que, si se comparan los estudios de sintaxis con los de fonética o morfología históricas, son obvios tanto el retraso metodológico como la escasez de los trabajos sintácticos.

[Ridruejo (1992, pg. 587), énfasis nuestro]

Me atrevo a decir que la sintaxis ha pasado de ser el patito feo de la lingüística histórica, romance y general, con una escasez notoria de estudios y estudiosos hace cincuenta años —si la comparamos con la fonología, y en buena parte, la morfología históricas— a ser el cisne de las subdisciplinas diacrónicas hoy en día [...]. Por décadas se dio una escisión tajante entre lingüistas sincrónicos y lingüistas diacrónicos, que conllevó el retraimiento de los estudios diacrónicos, muy especialmente los de sintaxis histórica. Subyacen a esta escisión varias razones teóricas.

[Company Company (2005, pgs. 144–146), énfasis nuestro]

En las últimas décadas se han puesto en marcha numerosas líneas de investigación que pretenden hacer una transición entre los estudios de corte tradicional de la gramática del español (notablemente, las gramáticas de Andrés Bello, Salvador Fernández Ramírez, o la misma Real Academia Española) y los estudios más actuales. Por lo general, esos esfuerzos han adoptado la forma

de monografías, manuales, tesis doctorales inéditas, artículos publicados e incontables actas de congresos.

Puede afirmarse, no obstante, que hay un punto de inflexión en dicha tendencia con la aparición de la *Gramática Descriptiva de la Lengua Española* (Bosque & Demonte, 1999) y, especialmente, con la *Nueva Gramática de la Lengua Española* (RAE-ASALE, 2009), obras en las que encontramos capítulos dedicados a fenómenos sintácticos en los que se recogen diferentes casos de variación. Es interesante observar, aun así, que la expresión “variación sintáctica” solo se encuentra en tres epígrafes de la NGLE (§§ 34.11.d, 34.11.e y 41.12o), lo cual da a entender la escasa popularidad que ese tipo de variación ha experimentado. Esfuerzos similares a los de RAE-ASALE (2009) se han llevado a cabo para otras lenguas, como el catalán (con la *Gramàtica del Català Contemporani*, 2002), el italiano (con la *Grande Grammatica Italiana di Consultazione*, 1991), el inglés (con la *Cambridge Grammar of the English Language*, 2002), y el portugués (con la *Gramática do Português Contemporâneo*, 2014).

1.2 Otras herramientas actuales

La ausencia de estudios de variación sintáctica en el caso del español contrasta con la situación de otras lenguas cercanas, que sí poseen atlas (o bases de datos) sintácticos. Hay diferentes ejemplos de ello:

1. El *Dynamic Syntactic Atlas of the Dutch dialects* [DynaSAND] <http://www.meertens.knaw.nl/sand/zoeken/index.php>
2. El *Syntax-oriented Corpus of Portuguese Dialects* [CORDIAL-SIN] <http://www.clul.ul.pt/en/resources/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects>
3. El *Atlas Linguístico-Etnográfico de Portugal e da Galiza* [ALEPG] <http://www.clul.ul.pt/en/resources/205-linguistic-and-ethnographic-atlas-of-portugal-and-galicia-alepg?showall=1>
4. El *Atlante Sintattico d'Italia* [ASIt] <http://asit.maldura.unipd.it/>
5. La *Base de Datos de la Sintaxis Vasca* [BASYQUE] <http://ixa2.si.ehu.es/atlas2/index.php?lang=es>

Junto a estos atlas y bases de datos, otras herramientas en línea han sido desarrolladas en los

últimos años. Lo más relevante es que todas ellas incorporan fenomenología sintáctica:

1. El *World Atlas of Linguistic Structures* [WALS] <http://wals.info/>
2. El proyecto *TERRALING* <http://www.terraling.com/>
3. El *Syntactic Structures of the World's Languages* [SSWL] <http://sswl.railsplayground.net/>
4. El proyecto *Symila* <http://blogs.univ-tlse2.fr/symila/en/>

Como hemos dicho, no existe, al menos de forma exclusiva, ninguna herramienta análoga en el caso del español. Visto con cierta perspectiva, dicha ausencia probablemente se debe a la conjunción de diferentes factores. Uno de ellos es la falta de herramientas formales suficientes para reflejar adecuadamente la variación sintáctica que presentan no solo lenguas tipológicamente diferenciadas (p.ej., español / cingalés), sino también lenguas cercanas (p.ej., español / catalán) o geoelectos de estas (p.ej., español de Santander / español de Buenos Aires). Como veremos en la siguiente sección, tales herramientas existen dentro de marcos teóricos formales como el de la Gramática Generativa.

1.3 La variación sintáctica: estado de la cuestión

En el contexto que estamos comentando, es importante destacar la contribución a la variación morfo-sintáctica de los enfoques formales, especialmente desde la llamada TEORÍA DE PRINCIPIOS Y PARÁMETROS (TPP; cf. Chomsky (1981, 1986)), en la que el lenguaje es concebido como una facultad biológica que consta de un estado inicial (E_I) y un estado final (E_F), que se corresponden con lo que se ha dado en llamar Gramática Universal (GU) y Gramática Particular (GP) respectivamente. La gran aportación del modelo TPP fue, precisamente, intentar reconciliar la supuesta inmutabilidad de la GU (los “principios”) con la obvia variabilidad de las diferentes GPs (los “parámetros”). En particular, la TPP permitió aliviar la tensión existente entre una descripción robusta de la variación lingüística (que pretendía caracterizar la variedad de cada lengua) y la explicación de la adquisición lingüística (un proceso aparentemente simple y sin instrucción explícita, incompatible con la proliferación de reglas que debían ser memorizadas por el niño que adquiere su lengua) mediante la sustitución de los sistemas de reglas

por principios universales que se fijaban mediante la experiencia lingüística. Desde tal perspectiva, un niño lo único que tendría que hacer es fijar un principio como el de (1), famoso en la bibliografía, de manera positiva (en el caso del inglés o el francés) o negativa (en el caso del español o el vasco):

- (1) Parámetro del Sujeto Nulo (PSN)
El sujeto de una oración debe manifestarse fonéticamente

Como puede verse, (1) únicamente habla de la posibilidad de que un sujeto se exprese fonéticamente (e.g., *los estudiantes han aprobado*) o no (e.g., \emptyset *han aprobado*). Sin embargo, pronto se observó que las lenguas que fijaban negativamente (1), desplegaban otras características (un “racimo” de propiedades; cf. Chomsky (1981, 1986)):

- (2) Efectos de racimo del PSN
 - a. Sujetos nulos:
 \emptyset *han aprobado*
 - b. Sujetos postverbiales:
Han aprobado ellos
 - c. Extracción larga de sujeto:
¿Quién dices que vino?
 - d. Pronombres expletivos nulos:
 \emptyset *Llueve*

Este tipo de correlaciones guardan ciertas similitudes con los “universales implicativos” propuestos por Joseph Greenberg (cf. Greenberg (1963); Mairal & Gil (2006), y referencias allí citadas), que tienen el formato “Si una lengua L tiene un rasgo x , entonces L tiene un rasgo y ”. En (3) ofrecemos una muestra de ese tipo de universales:

- (3) Universales implicativos
 - a. **Universal 5.** Si una lengua tiene el orden dominante SOV y el complemento genitivo aparece después del sustantivo, entonces el adjetivo también aparece después del sustantivo
 - b. **Universal 7.** Si en una lengua con orden dominante SOV, no hay orden alternativo básico, o solo OSV es una alternativa, entonces todos los modificadores adverbiales del verbo también preceden al verbo
 - c. **Universal 13.** Si el objeto nominal siempre precede al verbo, entonces las formas verbales subordinadas al verbo principal también lo preceden [tomados de Greenberg (1963, pgs. 79, 80, 64), traducción nuestra]

Acabamos de ver un parámetro cuya fijación tiene consecuencias para otros rincones de la gramática (los efectos de racimo), pero no hemos dicho mucho sobre dónde ni cómo se codifican los “puntos de variación”. Borer (1984) fue el primer trabajo en el que se planteó la pregunta de dónde se encuentran los parámetros —es decir, en qué componente de la gramática. Tal y como acabamos de ver, la TPP concibe los parámetros como principios moldeados mediante la experiencia, por lo que podríamos pensar que aquellos están en cualquiera de los módulos que posee una gramática: léxico, morfología, fonética, sintaxis, etc. Junto con Borer (1984), autores como Fukui (1986), Fukui & Speas (1986), Kayne (2000, 2005), Ouhalla (1991) y Webelhuth (1992) replantearon la propuesta de Chomsky (1986) al suponer que los parámetros se encontraban en el léxico. Borer (1984) expuso esta idea de manera clara al afirmar que “no hay elecciones de cada lengua específica con respecto a los principios y procesos universales. En vez de eso, la variación entre lenguas debería estar restringida a las propiedades idiosincrásicas de las unidades léxicas” (Borer (1984, pg. 2), traducción nuestra). Adoptando el término acuñado por Baker (2008), esta hipótesis se ha conocido como CONJETURA BORER-CHOMSKY (CBC):

- (4) Conjetura Borer-Chomsky (CBC)
 Todos los parámetros de variación son atribuibles a diferencias en los rasgos de ítems particulares (p.ej., núcleos funcionales) en el léxico [tomado de Baker (2008, pg. 353), traducción nuestra]

La CBC (y sus diversas manifestaciones) marcó un punto de inflexión para el estudio de la variación lingüística, especialmente al permitir distinguir dos tipos de parámetros: (i) MICROPARÁMETROS y (ii) MACROPARÁMETROS. El PSN es un ejemplo clásico de macroparámetro: se trata de un parámetro anclado a la sintaxis con consecuencias a gran escala una vez fijado y suele servir para distinguir lenguas tipológicamente distantes. Los microparámetros, por su lado, son parámetros compatibles con la CBC, con consecuencias limitadas a aquellas propiedades que puedan codificarse en las unidades léxicas, y con efectos visibles en lenguas cercanas (o geolectos de estas, lo cual es particularmente relevante para el estudio desarrollado en el ASinEs). La caracterización que de unos y otros ofrece Baker (2008) es esta:

La visión microparamétrica estándar es que las diferencias primitivas científicamente significativas entre lenguas [o

dialectos] son siempre diferencias pequeñas, típicamente asociadas a (como mucho) unas cuantas construcciones relacionadas [...] Las grandes diferencias entre lenguas siempre se reducen a muchas de las diferencias pequeñas [...] Por el contrario, la visión macroparamétrica es que hay al menos unos pocos parámetros (no compuestos) que definen tipológicamente diferentes tipos de lenguas.

[Baker (2008, pgs. 255–256), traducción nuestra]

Las principales asimetrías entre el punto de vista macroparamétrico y microparamétrico pueden resumirse como en el cuadro 1.

	Macro Parámetros	Micro Parámetros
DÓNDE	gramática (o sintaxis)	léxico
CÓMO	efectos masivos lenguas no relacionadas	efectos limitados lenguas relacionadas
QUÉ	filogenéticamente	filogenéticamente

Cuadro 1: Principales asimetrías entre el punto de vista macroparamétrico y microparamétrico

Es importante observar, para los propósitos del presente proyecto, que existen determinados fenómenos sintácticos del español a los que sí se ha prestado atención en los estudios geolectales (cf. Bosque (1999); Brucart (1994); Demonthe (2000); Demonte & Fernández-Soriano (2005); García Mouton (1994); Gómez Torrego (1999); Fernández-Ordóñez (1993, 1999); Ordóñez & Olarrea (2006)). Son los siguientes:

- (5) a. Dequeísmo:
Me dijo de que vendría tarde
 b. Queísmo:
La idea que no la volveré a ver más
 c. Relativas enfáticas:
No tienes idea de las cosas que dice
 d. Perífrasis (focales) de relativo:
Es en Boston que lo vi
 e. Leísmo:
Le criticaron duramente (de persona)
Ese libro, no le he leído (de cosa)
 f. Laísmo:
La dije la verdad a María
 g. Loísmo:
Lo di el libro, a Juan
 h. Duplicación pronominal:
La toqué a la sonata
 i. Distinción indicativo/subjuntivo:
No sé qué te diga

Aunque, como acabamos de decir, los fenómenos de (5) tienden a presentarse como sintácticos, creemos que, salvo los casos de (5-c) y (5-d), los demás podrían considerarse morfológicos (sobre todo si se adopta el marco teórico de la “Morfología Distribuida”; cf. Halle & Marantz (1993)). Esto plantea la cuestión de si existen fenómenos genuinamente sintácticos que manifiesten variación en el español (y en cualquier otra lengua; cf. Picallo (2014)). Sea cual sea la respuesta a esa pregunta, parece evidente que existe una asimetría entre los estudios sintácticos y los puramente morfo-fonéticos de la variación geolectal. Creemos, como decíamos hace un momento, que ello es el resultado de más de un factor. Además de la ausencia de un marco teórico que permita analizar las sutilidades de algunos casos de variación, la influencia del estructuralismo también dificultó que los estudios de variación aborasen datos sintácticos. En un trabajo reciente, Noam Chomsky observa lo siguiente:

La publicación de lo que fue la fundación de la lingüística estructuralista americana, *Métodos de lingüística estructural*, de Zellig Harris, se llamó “métodos” porque parecía haber poco que decir sobre el lenguaje más allá de los métodos que había para reducir los datos de lenguas que variaban sin límite a una forma organizada. El estructuralismo europeo fue esencialmente idéntico. La introducción clásica de Nikolai Trubezkoy al análisis fonológico tenía una concepción similar. De manera más general, las investigaciones estructurales se centraron casi exclusivamente en la fonología y la morfología, las áreas en las que las lenguas parecen diferir ampliamente y de manera compleja.

[apud Chomsky (2008, pgs. 2-3),
traducción nuestra]

Asumiendo que el primero de estos factores ha sido corregido (en la actualidad sí contamos con una teoría sintáctica que nos permite desarrollar análisis detallados), el proyecto ASinEs pretende paliar el vacío existente en los estudios de sintaxis geolectal del español mediante la elaboración de un atlas que tenga una base de datos asociada.

La ausencia a la que nos estamos refiriendo es doblemente sorprendente: por un lado, porque muchas otras lenguas poseen este tipo de herramientas y, por el otro, porque las peculiaridades del español, tanto cuantitativas (es la segunda lengua más hablada del mundo, solamente

superada por el chino mandarín, según los datos de Ethnologue²) como cualitativas (actualmente, está en contacto con lenguas románicas, germánicas, amerindias, así como con lenguas como el vasco) lo convierten en una fuente de interés tipológico para estudios de tipo geolectal. Por ello mismo, nuestro proyecto tiene un componente transversal, de “espectro amplio”, que lo hace interesante para diferentes usuarios: desde geolingüistas hasta sintactistas teóricos, pasando por sociolingüistas, tipólogos y estudiantes / profesores de español.

2 Objetivos del proyecto ASinEs

El proyecto ASinEs pretende desarrollar un estudio detallado y pionero, así como una caracterización precisa, de la variación sintáctica manifestada en los diferentes geolectos del español, una lengua hablada por aproximadamente 420 millones de hablantes nativos (460 millones, si tenemos en cuenta a los hablantes que tienen el español como lengua segunda), mayoritariamente en la Península Ibérica y el continente americano.

Una iniciativa de tales características no tiene precedentes en el campo (no existe ningún estudio sistemático de la variación sintáctica del español), y mucho menos con una herramienta en línea: un atlas que pretende incorporar una amplia base de datos, creada a partir de las gramáticas de referencia del español (GDLE 1999 y NGLLE 2009). De manera más general, nuestro proyecto contribuye no solo a la creación de una hoja de ruta detallada de la sintaxis del español (un objetivo con consecuencias muy provechosas para investigadores con diferentes posicionamientos teóricos, como hemos apuntado), sino también a la comprensión de un fenómeno complejo (la facultad del lenguaje y sus diferentes manifestaciones) que posee un amplio interés para otras áreas, de tipo cognitivo, histórico y sociológico (p.ej., bilingüismo, biodiversidad, antropología lingüística, diglosia, conflictos sociopolíticos, *code-switching*, *Sprachbund*, desarrollo del lenguaje, lenguas heredadas, sistemas de comunicación, psicolingüística, etc.).

El proyecto se articula alrededor de la creación de dos herramientas compactadas que ejercen de eje vertebrador:

- **Una base de datos / corpus de fenómenos y construcciones sintácticas del español.** Los datos son extraídos, principalmente, de las gramáticas de referencia del español, la GDLE (1999) y la NGLLE (2009),

²<https://www.ethnologue.com>

aunque también se tienen en cuenta monografías, tesis doctorales, artículos publicados y otros materiales. En fases posteriores del proyecto, se incorporará información extraída de trabajo de campo (encuestas, entrevistas, etc.)

- Un atlas interactivo en línea del español (ASinEs), que se encarga de reflejar la variación sintáctica de las variantes del español, tanto americanas como europeas.

Este proyecto es original al menos por dos razones. Por un lado, a nivel teórico, el desarrollo de ASinEs conlleva, por primera vez, una investigación sistemática sobre la sintaxis geolectal del español, aprovechando para ello una base teórica y tipológica robusta, y también incorporando el reciente trabajo descriptivo de la GDLE (1999) y la NGLE (2009). Además, nuestra iniciativa aborda de manera sistemática aspectos que atañen al contacto de lenguas y la transición geolectal, incorporando a tal efecto tanto las estrategias tradicionales de investigación geolectal (la bibliografía disponible y los atlas existentes) como los mecanismos más actuales (corpus informáticos, bases de datos y atlas en línea, como el WALS, el BASYQUE, el SSWL o el TERRALING).

Por otro lado, a nivel metodológico, la investigación llevada a cabo tomando el español como objeto de estudio proporciona dos herramientas de trabajo que no tienen precedentes: una base de datos de construcciones con variación sintáctica y un atlas interactivo capaz de geolocalizar los puntos de variación. En este sentido, es importante volver a destacar el estatus especial del español en términos de cobertura empírica y geográfica, número de hablantes, y número de lenguas con las que interactúa (francés, catalán, gallego, vasco, inglés, árabe, rumano, etc.).

En resumen, tanto teórica como metodológicamente, el actual proyecto se adentra en territorio inexplorado. Los resultados de esta experiencia proporcionarán, sin duda, una visión novedosa y estimulante que servirá de referente para proyectos similares en el dominio de la lingüística, así como para otros ámbitos en los que el trabajo interdisciplinar es necesario. Y, naturalmente, el ASinEs como tal se convertirá en una herramienta para investigadores no solo de la variación sintáctica del español, sino también de la variación sintáctica de otras lenguas con las que esta esté en contacto (tanto románicas como no románicas).

3 Funcionamiento de la aplicación

La aplicación ASinEs (cuya interfaz de entrada puede verse en la Figura 2) utiliza la tecnología de Google Maps para geolocalizar los fenómenos sujetos a estudio; es decir, las construcciones sintácticas que presenten variación geolectal.



Figura 2: Interfaz principal del ASinEs.

La base de datos se compone de un sistema de fichas que contienen la siguiente información:

- (6) Información fichas ASinEs
Ejemplo: La dije que no era verdad
Fuente: GDLE (§§ 23.4., 11.5), NGLE (§§ 34.1., 30.2., 4.7)
Lengua: español
Dialecto: español de Castilla
Fenómeno: laísmo
Construcción: construcción de doble objeto
Elem. gram.: pronombre, determinante, etc.
relacionado con: leísmo, loísmo
¿Se encuentra en otras lenguas? No
¿Se encuentra en otros estadios del español? No
Bibliografía: Romero (1997)
Geolocalización: puntos / áreas (predefinidas) / áreas (dibujo libre)
Sexo del informante (si es relevante)
Edad del informante (si es relevante)
Formación del informante (si es relevante)
Archivo de audio / vídeo
Comentarios:
Autor:
Fecha de creación:

A continuación, en la Figura 3, ofrecemos un ejemplo concreto de las fichas del ASinEs.

Como puede verse, el motor de búsqueda del ASinEs incluye análisis morfológicos y sintácticos. El sistema permite utilizar cualesquiera de los campos para realizar búsquedas, ya sean simples (basadas en un parámetro de las fichas) o

asines1056

Vista Editar Generate PDF Log

+ Clone content

Enviado por **acerrudo** el Mié, 06/17/2015 - 16:26

No sé de qué secreto me estás hablando.

No sé de qué secreto me estás
NEG know.1SG of what secret CL.DAT.1SG are.2SG

'I don't know what secret you are talking about.'

REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [en línea]. *Corpus de referen*

Dialecto:
Área rioplatense Europeo

Distribución geográfica:
"En algunas áreas, entre ellas, la rioplatense y la europea." (NGLE: §22.14a)

Construcción:
Oración interrogativa indirecta

Detalle 2:
Alternancia {cuál -que} como determinantes interrogativos

Elementos gramaticales:
Determinante Determinante interrogativo Modalidad interrogativa

Etiquetado morfológico:
No (ADV) sé (V) de (P) qué (D.INT) secreto (N) me (CL) estás (V.AUX) habla

¿Presente en otras lenguas?:
Inglés

Figura 3: Ejemplo de ficha del ASinEs.

cruzadas (combinando más de un parámetro). De manera adicional, se incorpora un etiquetado morfológico y sintáctico, como se indica en (7):

- (7) Ejemplo: *María dijo la verdad a Pedro*
 ETIQUETADO MORFOLÓGICO
María(N) dijo(V) la(D) verdad(V) a(P) Pedro(N)
 ETIQUETADO SINTÁCTICO
 [(O) María [(SV) dijo [(SN) la verdad] [(SP) a Pedro]]]

Un etiquetado de doble nivel, como el de (7), ofrece la posibilidad de realizar búsquedas de símbolos terminales (cadenas markovianas, generables por máquinas de estados finitos) o secuencias que combinen símbolos terminales con símbolos no terminales. Ambos tipos de búsquedas se ilustran en (8):

- (8) Búsqueda 1: V + N > dice cosas, etc.
 Búsqueda 2: V + SN > dice [SN esas cosas], etc.

La interfaz de búsqueda asociada a la base de datos del ASinEs aparece en la Figura 4.

El tipo de sistema que estamos implementando permite, asimismo, desarrollar un analizador (*parser*), capaz de etiquetar automáticamente. Este analizador puede utilizar información disponible en las redes sociales (que, por defecto, contienen información geográfica) para complementar la información extraída de las gramática de referencia. Además, el sistema de etiquetado de (7) empleará una sintaxis que puede ser utili-

Figura 4: Ejemplo de menú de búsqueda del ASinEs.

zada para generar diagramas arbóreos³, como los que se muestran en la Figura 5.

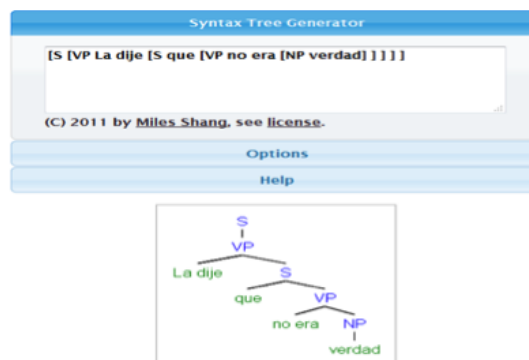


Figura 5: Ejemplo del analizador sintáctico.

Por su parte, el sistema de geolocalización permite asociar áreas geolectales concretas a las variantes del español relevantes, como se ve en las Figuras 6, 7 y 8.



Figura 6: Geolocalización (área) de Colombia.

Este tipo de información, codificable con puntos o áreas, está presente en muchas de las redes sociales, lo cual abre considerablemente el abanico de fuentes de las que extraer información para caracterizar tal o cual construcción sintáctica.

³Como, por ejemplo, <http://mshang.ca/syntaxtree/>



Figura 7: Geolocalización (área) de Cataluña.



Figura 8: Geolocalización (área) de Castilla-La Mancha.

Entre estas redes se encuentran Instagram (utilizada sobre todo por jóvenes para comentar fotografías), Facebook (utilizada por usuarios de todas las edades, para conectar con amigos y conocidos), LinkedIn (utilizada por adultos por motivos profesionales), Tripadvisor (utilizada para compartir opiniones sobre restaurantes, hoteles, etc.) o Twitter (utilizada sobre todo por usuarios adultos con fines laborales o socioculturales). Toda esa información podría incorporarse al ASinEs en el futuro para complementar los datos conseguidos a través de las gramáticas de referencia, libros y manuales, tesis doctorales y artículos publicados, o información recopilada mediante trabajo de campo.

En resumen, el proyecto ASinEs ofrece una herramienta versátil para la investigación de la variación sintáctica de los geolectos del español, diseñada alrededor de una base de datos y un atlas que proporcionan diversos parámetros de búsqueda.

4 Conclusiones

El objetivo de este artículo ha sido el de presentar el Proyecto ASinEs, una aplicación que desarrolla un atlas sintáctico de la variación de los geolectos del español. Para ello, hemos revisado los antecedentes existentes (en los estudios geolectales de corte tradicional) y cuáles han sido

los cambios teóricos que han permitido abordar el estudio de la variación sintáctica de manera realista.

El presente proyecto tiene, como se ha visto, objetivos ambiciosos, y pretende convertirse en una herramienta para los investigadores y estudiantes de variación sintáctica tanto del español como de otras lenguas relacionadas. En una primera fase (iniciada en enero de 2015) nos hemos centrado en la confección de una base de datos que recoja los puntos de variación presentes en las gramáticas de referencia del español, la *Gramática Descriptiva de la Lengua Española* (1999) y la *Nueva Gramática de la Lengua Española* (2009). En fases posteriores, pretendemos incorporar información de las siguientes fuentes:

- Tesis, monografías y artículos publicados
- Trabajo de campo (entrevistas, cuestionarios, etc.)
- Corpus que codifiquen información sintáctica relevante para el ASinEs (COSER, BASYQUE, etc.)
- Redes sociales (Twitter, Facebook, etc.)

El proyecto que presentamos en estas páginas es, en definitiva, de largo recorrido y puede ampliarse si se desarrollan las líneas de colaboración internacionales que existen en el Proyecto Edisyn⁴.

Agradecimientos

Los autores de este artículo queremos dejar constancia de las diferentes fuentes de financiación que han permitido trabajar en el desarrollo del Atlas Sintáctico del Español (ASinEs). En primer lugar, todos agradecemos a la Fundación BBVA la concesión de una *Ayuda a Investigadores, innovadores y creadores culturales 2014*, que ha permitido poner en marcha el proyecto. En segundo lugar, agradecemos también las ayudas que provienen de la Generalitat de Catalunya (2014SGR-1013 y 2014SGR-1511) y el Ministerio de Economía y Competitividad (FFI2014-56968-C4-2-P, FFI22011-29440-C03-03, FFI2012-31415 y FFI2014-56968-C4-4-P). Gracias, por último, a los útiles y acertados comentarios de dos revisores anónimos. Si hay algún error, es nuestro.

⁴http://www.dialectsyntax.org/wiki/Main_Page

Referencias

- Alvar, Manuel. 1996a. *Manual de dialectología hispánica. el español de américa*. Barcelona: Ariel.
- Alvar, Manuel. 1996b. *Manual de dialectología hispánica. el español de españa*. Barcelona: Ariel.
- Baker, Mark C. 2008. *The syntax of agreement and concord*. Cambridge University Press.
- Borer, Hagit. 1984. *Parametric syntax: Case studies in semitic and romance languages*. Dordrecht: Foris Publications.
- Bosque, Ignacio. 1999. Sobre la estructura sintáctica de una construcción focalizadora. En *En Homenaje al profesor Ambrosio Rabanales*, 207–231. BFUCh XXXVII.
- Bosque, Ignacio & Violeta Demonte (eds.). 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe.
- Brucart, José María. 1994. Syntactic variation and gramatical primitives in generative grammar. En A. Briz y M. Pérez-Saldanya (ed.), *Lynx. Categories and functions. A monographic series in Linguistics and world perception*, 145–176. Publicacions de la Universitat de València/University of Minnesota.
- Campbell, Lyle. 2001. The history of linguistics. En Mark Aronoff & Janie Rees-Miller (eds.), *The Handbook of Linguistics*, 81–104. London: Blackwell Publishers.
- Chambers, Jack K. & Natalie Schilling-Estes (eds.). 2013. *The handbook of language variation and change*, vol. 129. John Wiley & Sons.
- Chambers, Jack K. & Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1981. *Lectures on government and binding: The pisa lectures*. Dordrecht, Holland: Foris Publications.
- Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam. 2008. The biolinguistic program: Where does it stand today? Ms. MIT.
- Company Company, Concepción. 2005. Una paradoja de la lingüística histórica romance: el florecimiento de la sintaxis histórica románica. *La Crónica* 34(1). 144–163.
- Company Company, Concepción (ed.). 2006. *Sintaxis histórica de la lengua española. primera parte: La frase verbal*. México: Fondo de Cultura Económica-Universidad Nacional Autónoma de México.
- Company Company, Concepción (ed.). 2009. *Sintaxis histórica de la lengua española. segunda parte: La frase nominal*. México: Fondo de Cultura Económica-Universidad Nacional Autónoma de México.
- Company Company, Concepción (ed.). 2014. *Sintaxis histórica de la lengua española. tercera parte: Adverbios, preposiciones y conjunciones. relaciones interoracionales*. México: Fondo de Cultura Económica-Universidad Nacional Autónoma de México.
- Demonte, Violeta. 2000. Gramática, variación y norma. una tipología. *Estudios Hispánicos (Revista de la Sociedad Coreana de Hispanistas)* 17(12). 3–49.
- Demonte, Violeta & Olga Fernández-Soriano. 2005. Features in comp and syntactic variation: the case of '(de)queísmo' in spanish. *Lingua* 115(8). 1063–1082.
- Fernández-Ordóñez, Inés. 1993. Leísmo, laísmo y loísmo: estado de la cuestión. En O. Fernández Soriano (ed.), *Los pronombres átonos*, Madrid: Taurus.
- Fernández-Ordóñez, Inés. 1999. Leísmo, laísmo y loísmo. En I. Bosque & V. Demonte (eds.), *Gramática descriptiva de la lengua española*, Madrid: Espasa Calpe.
- Fernández-Ordóñez, Inés. 2011. La lengua de castilla y la formación del español. Discurso leído el 13 de febrero de 2011 en su recepción pública por la Excma. Sra. D.^a Inés Fernández-Ordóñez y contestación del Excmo. Sr. D. José Antonio Pascual. Madrid.
- Fernández-Ordóñez, Inés. 2014. Los dialectos del español de España. En Javier Gutiérrez Re-xach (ed.), *Enciclopedia lingüística hispánica*, Routledge. En prensa.
- Fukui, Naoki. 1986. *A theory of category projection and its applications*: MIT. Tesis Doctoral.
- Fukui, Naoki & Margaret Speas. 1986. Specifiers and projections. *MIT Working Papers in Linguistics* 8. 128–172.
- García Mouton, Pilar. 1994. *Lenguas y dialectos de españa*. Madrid: Arco/Libros.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. En Joseph H. Greenberg (ed.), *Universals of Human Language*, 73–113. Cambridge, Mass: MIT Press.

- Gómez Torrego, Leonardo. 1999. La variación en las subordinadas sustantivas: dequeísmo y queísmo. En Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, Madrid: Espasa Calpe.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. En K. Hale & S.J. Keyser (eds.), *The view from Building 20: Essays in honour of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.
- Kany, Charles Emil. 1945. *American spanish syntax*. Chicago: The University of Chicago Press. [traducción esp. *Sintaxis hispanoamericana*. Madrid: Gredos. 1970].
- Kayne, Richard S. 2000. *Parameters and universals* Oxford studies in comparative syntax. Oxford University Press, USA.
- Kayne, Richard S. 2005. *Movement and silence* Oxford Studies in Comparative Syntax. Oxford University Press, USA.
- Labov, William. 1994. *Principles of linguistics change: Internal factors*, vol. 1. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistics change: Social factors*, vol. 2. Oxford: Blackwell.
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of north american english: Phonetics, phonology and sound change*. Berlin: Mouton/de Gruyter.
- Lapesa, Rafael. 2000. *Estudios de morfosintaxis histórica del español*. Madrid: Gredos.
- Mairal, Ricardo & Juana Gil. 2006. *Linguistic universals*. Cambridge: Cambridge University Press.
- Morillo-Velarde Pérez, Ramón. 1992. Un modelo de variación sintáctica dialectal: el demostrativo de realce en andaluz. En M. Ariza, R. Cano, J. M.a Mendoza & A. Narbona (eds.), *Actas del II Congreso Internacional de Historia de la Lengua Española*, vol. 2, 219–227. Madrid: Pabellón de España.
- Ordóñez, Francisco & Antxon Olarrea. 2006. Microvariation in caribbean/non caribbean spanish interrogatives. *Probus* 18(1). 59–96.
- Ouhalla, Jamal. 1991. *Functional categories and parametrization*. London: Routledge.
- Petyt, Keith Malcolm. 1980. *The study of dialect: An introduction to dialectology*. London: The language library.
- Picallo, M. Carme. 2014. *Linguistic variation in the minimalist framework*. Oxford: Oxford University Press.
- RAE-ASALE. 2009. *Nueva gramática de la lengua española*. Madrid: Espasa.
- Ridruejo, Emilio. 1992. Sintaxis histórica. En *Actas del I Congreso de la Lengua Española*, http://www.cvc.cervantes.es/obref/congresos/sevilla/unidad/ponenc_eridruejo.htm.
- Sánchez Lobato, Jesús. 1994. El español en américa. En *Problemas y métodos en la enseñanza del español como lengua extranjera. Actas del IV Congreso Internacional de ASE-LE*, 553–570. Madrid.
- Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. Oxford: Oxford University Press.