

DEL ATLAS LINGÜÍSTICO TRADICIONAL AL CORPUS GEOLINGÜÍSTICO DIGITAL: DISEÑO DE UN PROYECTO¹

CAROLINA JULIÀ LUNA

Universidad Nacional de Educación a Distancia

cjulialuna@flog.uned.es

ORCID: <https://orcid.org/0000-0001-5599-6675>

RESUMEN

El presente artículo muestra el modo en el que las investigaciones filológicas, gracias a la revolución digital, han aumentado el interés por representar los datos lingüísticos en mapas. Desde el punto de vista de la geolingüística esto ha generado, por un lado, el inicio de numerosos proyectos que tienen como fin la informatización de los atlas lingüísticos y, por el otro, el surgimiento de un número nada desdeñable de herramientas digitales que a pesar de alejarse del método de la geografía lingüística tradicional representan la información en mapas. En este contexto, nace la idea del proyecto CORPAT (*Corpus de los atlas lingüísticos*), un corpus geolectal cuyo objetivo principal es el de recopilar las respuestas de los atlas lingüísticos regionales del español para preservarlas y ofrecerlas a la comunidad científica y lingüística con el fin de que puedan ser estudiadas como complemento a los datos que ofrecen los corpus textuales y las obras lexicográficas digitales.

PALABRAS CLAVE: variación dialectal; humanidades digitales; atlas lingüísticos; corpus dialectales

FROM THE TRADITIONAL LINGUISTIC ATLAS TO THE DIGITAL GEOLINGUISTIC CORPUS: DESIGN OF A RESEARCH PROJECT

ABSTRACT

This article shows the way in which philological research has increased interest in representing linguistic data on maps due to the digital revolution. From a geolinguistics point of view this has generated, on the one hand, the beginning of numerous projects that aim to computerize traditional linguistic atlases and, on the other hand, the emergence of a not inconsiderable number of digital tools to represent the information on maps despite their moving away from the method of traditional linguistic geography. In this context, the idea of the *CORPAT (Corpus de los atlas lingüísticos)*, a geolectal corpus whose main objective is to compile the answers of the regional linguistic atlases of Spanish to preserve them and offer them to the scientific and linguistic community so that they can be studied as a complement to the data offered by textual corpora and digital lexicography, was born.

KEY WORDS: Dialectal variation; Digital Humanities; Linguistic atlas; Dialectal corpora

1. UN POCO DE HISTORIA

El interés por el estudio de la variación lingüística se acrecienta a principios del siglo XIX y va en aumento a lo largo de toda la centuria del ochocientos en paralelo con el desarrollo de la lingüística como disciplina científica. El punto de partida de este cambio en el enfoque de los estudios sobre las lenguas se debe a que se empieza «a realizar un trabajo empírico de observación de los datos» (Tusón 2003: 148) que deja de lado el análisis basado en especulaciones. En este marco nacen el comparatismo, la neogramática y la dialectología (Thibault 2007: 351). En el caso de los estudios

¹ Esta investigación ha sido posible gracias al apoyo del Departamento de Lengua Española y Lingüística General de la Universidad Nacional de Educación a Distancia y del Comissionat per Universitats i Recerca de la Generalitat de Catalunya (SGR 2017-1251).

dialectales románicos, puede resultar adecuado metodológicamente, según García Mouton (2007: 321), «situarse en la segunda mitad del siglo XIX, cuando ya los lingüistas habían descubierto el valor de los dialectos como reserva de materiales inestimables, porque iban a permitir ir mucho más allá en sus procesos de reconstrucción de lo que habían hecho a partir de las lenguas normativas».

El giro metodológico en los estudios sobre variación lingüística supone un cambio sustancial (Sanchis Guarner 1953: 7), ya que se dejaron de lado los antiguos análisis fundamentados en la interpretación de documentos escritos y empezaron a realizarse estudios basados en datos recogidos por los propios investigadores que equiparaban la investigación lingüística a la que se desarrollaba en las ciencias naturales (Malmberg 1983: 150). Entre los primeros trabajos empíricos, siguiendo a Pop (1950: 480), se puede mencionar, por un lado, el estudio de J. A. Schmeller (*Die Mundarten Bayerns*, 1821), que utiliza encuestas para estudiar las características de las hablas de Baviera; y, por el otro, el libro de Bernardino Biondelli (*Saggio sui dialetti gallo-italici*, 1853), en el que se describen por primera vez los dialectos italianos a partir del estudio de información recopilada por el autor.

Es en este contexto de efervescencia científica en el que nace el método de investigación dialectal por excelencia, la geografía lingüística o geolingüística (Coseriu 1977: 103; García Mouton 2007: 323), cuyo objetivo principal fue recoger información de los dialectos en mapas. Así, a finales de siglo XIX, el alemán Georg Wenker —desde la escuela neogramática y con el fin de demostrar la regularidad del cambio fonético— publica algunos mapas del que hoy se considera el primer atlas lingüístico, el *Sprachatlas von Nord-und Mitteldeutschland* (1881). Posteriormente, el método se establece gracias al investigador francés Jules Gilliéron, quien fundó las bases de la geografía lingüística a través de la publicación del *Atlas Linguistique de la France* (ALF, 1902-1910). De acuerdo con Lameli (2010: 567), debe tenerse en cuenta que el escenario en el que se crean los primeros mapas de palabras para representar la variación lingüística es muy favorable, pues coincide con un momento de esplendor de la ciencia cartográfica:

The production of linguistic maps became a relevant linguistic field in the nineteenth century, at which time we find an abrupt rise in the amount of cartographic work, making it possible to speak of the origin of the cartographic paradigm in linguistics. Right from the outset, atlases, not just maps, were seen as a productive tool in the pursuit of geolinguistic interests. The extent to which this was a part of a more general zeitgeist becomes clear when one reflects that geographers were also publishing the first thematic atlases at this time.

Desde que vio la luz el que se ha conocido como el primer atlas lingüístico hasta la actualidad han pasado 140 años, un tiempo en el que el método, lejos de quedar obsoleto, se ha mantenido y ha evolucionado de muy diversas formas (Lameli 2010; Hoch y Hayes 2010; Sousa en prensa; Julià en prensa). De entre todos los cambios que han afectado a la geolingüística, la aplicación de la tecnología a la creación de mapas es, sin duda, el más destacable, pues se ha pasado del mapa en papel al mapa digital, lo que ha supuesto considerables mejoras tanto en lo que se refiere a la preparación como a la consulta y explotación de los datos. Si se contrastan los primeros mapas publicados (figura 1) con los de más reciente creación (figura 2), se advierten

diferencias entre uno y otro; no obstante, también se aprecia que estas no suponen la pérdida de la esencia del atlas ni del mapa: el objetivo sigue siendo la representación cartográfica de datos lingüísticos con el fin de obtener una visión espacial de la extensión de los fenómenos de variación lingüística.

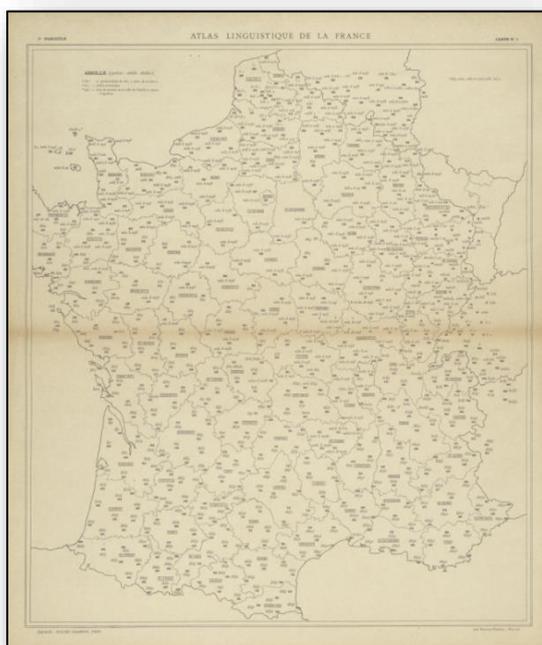


Figura 1. *Atlas Linguistique de la France* (ALF, 1902-1910, mapa 1, abeille)



Figura 2. *Atlas Dialectal de Madrid* (ADiM 2015, mapa 512, abejorro)

Desde un punto de vista histórico, la digitalización de los mapas lingüísticos supone, según Sousa (2017: 2), que la dialectología sea una de las primeras disciplinas humanísticas en incorporar el uso de la tecnología informática tanto en el procesamiento como en el análisis de datos lingüísticos. El inicio de la informatización de los atlas se sitúa a mediados de la década de 1970 (Hoch y Hayes 2010: 25) y sus primeros resultados son tres grandes obras europeas: «The three projects which stand out as pioneers are *Computer Developed Linguistic Atlas of England* (Viereck y Ramisch 1991-1997), *Atlas Linguarum Europae* (Alinei et al. 1983) and *Kleiner Deutscher Sprachatlas* (Veith et al. 1984-1999)» (Sousa 2017: 2). Además, no hay que obviar que «los primeros dialectómetros tuvieron que generar sus bases de datos partiendo de información extraída de atlas lingüísticos ya publicados» (Aurrekoetxea 2019: 23). En España, Manuel Alvar se unió al proceso de automatización e informatización de los atlas (Alvar y Verdejo 1978; Alvar y Nuño 1981) con el proyecto del *ALES* (*Atlas Lingüístico de Santander*) —al que hoy se conoce como *Atlas Lingüístico y Etnográfico de Cantabria* (*ALEcant*)—; sin embargo, el propio investigador describió el proceso como un intento de adaptación que no se produjo según lo esperado ni en tiempo ni en forma (Alvar 1991). Unos años más tarde, entre 1999 y 2001, David Heap revolucionó los estudios geolingüísticos de la península ibérica al recopilar y digitalizar los cuadernos de respuestas del *Atlas Lingüístico de la Península Ibérica* (*ALPI*) y

ponerlos en la red a disposición de la comunidad científica y de la sociedad (Heap 2002).

En la actualidad, gracias a la aplicación de las nuevas tecnologías —Bases de datos Espaciales (BDE) y Sistemas de Información Geográfica (SIG)— a la elaboración de mapas lingüísticos, disponemos de diversas muestras de atlas digitales finalizados o en proceso de elaboración. El presente artículo se divide en dos grandes bloques en los que se pretende, por un lado, ilustrar y examinar las posibilidades de explotación de esta geolingüística digital (§ 2) y, por el otro, presentar el proyecto *Corpus de los atlas lingüísticos (CORPAT)* (§ 3).

2. LA GEOGRAFÍA DIGITAL Y LA GEOLINGÜÍSTICA EN EL SIGLO XXI

La evolución de la tecnología en los últimos veinte años, los primeros del inicio del tercer milenio, ha provocado que la geolocalización —esto es, la ubicación geográfica de un objeto, lugar o persona— haya adquirido un papel fundamental en nuestras actividades cotidianas. Basta con pensar en que el uso de la aplicación Google Maps, a la que se puede acceder desde cualquier dispositivo de telefonía móvil para situar algo o a alguien en un mapa, se ha convertido en una herramienta de uso habitual, casi imprescindible, en nuestro quehacer diario. Realizar búsquedas de lugares, crear y guardar rutas de viaje, consultar el tráfico o compartir ubicaciones para poder localizar un punto de encuentro son solo algunos de los múltiples usos que permiten los mapas virtuales. Esto es consecuencia de las grandes ventajas que los SIG ofrecen al mundo y que, según algunos investigadores (Buzai 2004), han cambiado incluso la forma de pensar del ser humano en su día a día. Ello se debe, entre otros factores, al hecho de que la cartografía haya pasado a ser dinámica y podamos interactuar con los datos a través de su representación geográfica. La red ofrece numerosas posibilidades (programas descargables gratuitos, aplicaciones y recursos en línea) al usuario, sea experto o aficionado, para representar a su gusto informaciones en distintos tipos de mapas. [Google My Maps](#), [Gabmap](#), [Diatech](#), [OpenStreetMap](#), [ArcMap](#), [Carto](#), [Mapbox](#) o [QGIS](#) son una muestra de este tipo de programas que permiten crear un mapa a la carta. La mayoría son de carácter general, aunque también existen los especializados como Gabmap y Diatech, dos programas de código abierto y de uso en línea que permiten generar mapas para el estudio dialectométrico (Aurrekoetxea 2019: 20). Según Ruiz Tinoco (2018), uno de los más utilizados, por todo lo que permite hacer y por tratarse de software libre, es QGIS.

En el ámbito de las humanidades y, en especial, de la investigación lingüística, el uso de los SIG y las BDE —es decir, las que «almacenan entidades del mundo real que tienen componentes espaciales y no espaciales de forma estructurada y sin duplicidades» (Bonilla y Bernal 2020: 352)— ha repercutido en numerosos estudios y ha permitido crear un número nada despreciable de herramientas digitales que representan datos en el espacio geográfico (Lameli 2007; Hoch y Hayes 2010) con un gran potencial de análisis (Thomas 1980; Bosque *et al.* 2012).

Estas herramientas se pueden clasificar en dos grandes grupos: por un lado, las que siguen los postulados de la tradicional geografía lingüística y publican atlas en formato digital (§ 2.2); y, por el otro, las que metodológicamente se alejan del método geolingüístico en diferentes aspectos, pero que usan los mapas para recopilar y

representar los datos con el fin de dar cuenta de la variación lingüística a través del espacio (§ 2.3). Antes de adentrarnos en la descripción de estas herramientas digitales lingüístico-cartográficas, conviene detenerse en una cuestión conceptual relevante para comprender qué se entiende por *atlas lingüístico* y *mapa lingüístico* (§ 2.1).

2.1. Los conceptos de ‘atlas lingüístico’ y ‘mapa lingüístico’

En un primer acercamiento a la expresión *atlas lingüístico* es posible suponer, de forma acertada, que se trata de una publicación de carácter interdisciplinar en la que se vinculan la geografía y la lengua. Así, de forma simple, un atlas lingüístico se puede definir como un conjunto de mapas que representan información sobre las lenguas. Sin embargo, la designación *atlas lingüístico* puede usarse para hacer referencia a dos ideas distintas en función de los datos que ofrecen los mapas y del método empleado para recopilarlos (Lameli 2010: 567-568; Sousa en prensa). Por una parte, un atlas lingüístico puede ser una compilación de mapas en los que se indica el espacio que ocupan las lenguas. Esta acepción, de carácter general, es la que se refiere, según Sousa (en prensa), a los denominados atlas etnolingüísticos que fueron creados en el siglo xvi. Estos constituyen, para Coseriu (1977: 109-110), el producto de los estudios de la geografía de las lenguas (Orozco 1864). Una muestra de este tipo de atlas puede ser el [Atlas interactivo UNESCO de las lenguas del mundo en peligro](#) (Moseley 2010) en el que se marcan los lugares del mundo en los que existe una lengua en peligro de extinción y se aportan, además, datos sobre ella.

Por otra parte, la segunda acepción de *atlas lingüístico*, la más extendida desde punto de vista especializado y científico (Coseriu 1977), surge a finales del siglo xix y se concibe como un conjunto de mapas agrupados que reúnen «grandes cantidades de materiales recogidos de un modo riguroso según unas pautas previamente pactadas, y todo en una sincronía relativa» (García Mouton 2007: 324). En este caso, según Sousa (en prensa), estos son los atlas dialectales, ya que se sirven del mapa geográfico como herramienta para representar datos lingüísticos con el fin de poder examinar y determinar a partir de ellos las características de los dialectos. Además, cabe tener en cuenta que en la definición de este tipo de atlas es relevante la procedencia de los datos que en ellos se representan, pues, como se indica en la vigésimo tercera edición del *DLE* (2014) —que documenta por primera vez en su historia este sublema poliléxico—, un *atlas lingüístico* es el «conjunto de mapas en que se presentan datos lingüísticos procedentes de encuestas» (s. v. *atlas*). Sin embargo, como veremos en el epígrafe 2.3, en la última década se han publicado diversas herramientas informáticas a las que se ha etiquetado y designado con la expresión *atlas lingüístico* a pesar de no cumplir con estas propiedades. La diferencia entre la primera acepción y la segunda es la que originó que algunos investigadores no estuvieran de acuerdo en referirse a este método de estudio lingüístico como *geografía lingüística* (también, de forma más reciente, *geolingüística*) y creyeran que era mucho más acertado designarla, entre otras formas, *lingüística geográfica*² (Dauzat 1944 [1922]: 5).

² Además de *geografía lingüística* y *lingüística geográfica*, otros investigadores se han referido a este método como *geografía dialectal*, *dialectología regional* o *dialectología tradicional*, entre otras (Walters 1990-1992). Además, según Alvar (1951 *apud* Sanchis Guarner 1953: 9), si la geografía lingüística se proyecta desde una perspectiva diacrónica, podemos referirnos a esta metodología como *geología*

Los atlas lingüísticos tradicionales se componen de mapas lingüísticos que se caracterizan porque «[they] combine topographic with linguistic information and represent the areal distribution of linguistic variables» (Limper, Pheiff y Williams 2020: 3747). De acuerdo con Girth (2010), desde la perspectiva de la semiótica de Charles Peirce, el mapa del atlas lingüístico puede interpretarse según la estructura triádica del signo lingüístico en la que el conocimiento se entiende como un proceso de significación (Koval 2017):

The sign consists of a representamen, object, and interpretant. In the case of the cartographic sign, the representamen is its visual (and occasionally auditory) form. The object is the topographic and linguistic reality mirrored by it. Finally, the linguistic map is only an indirect interpretation of reality via the interpretant, the mental picture evoked in the map user. With this model, the linguistic map can be described on multiple levels: structural, thematic, and cognitive. Girth (2010, pp. 99-100) formulates these levels in terms of questions, which turn out to be distinct steps in the production of maps (Limper, Pheiff y Williams 2020: 3747).

La publicación digital y la aplicación de los SIG, como se explica en los siguientes epígrafes, benefician su interpretación desde esta y otras perspectivas.

2.2. El atlas lingüístico tradicional en formato digital

En la actualidad, existe una amplia variedad de instrumentos digitales para el estudio de la variación diatópica, la mayoría de los cuales se basa en la recopilación y organización en bases de datos de materiales escritos u orales; muchos de ellos, además, representan en mapas los fenómenos lingüísticos. Desde la perspectiva y el concepto tradicional de 'atlas lingüístico' —es decir, el que se corresponde con el que Jules Gilliéron ideó y materializó sobre Francia a principios del siglo xx y que sigue la metodología de recogida y representación de datos de la geolingüística—, pueden diferenciarse tres grandes grupos de proyectos³: los que tienen como finalidad la reproducción digital, a modo de facsímil, de los atlas tradicionales que se publicaron inicialmente en papel; los que ofrecen los mapas digitalizados y, además, brindan la posibilidad de consultar bases de datos que permiten realizar búsquedas por la información que contienen; y, finalmente, los que han diseñado y publicado un atlas lingüístico directamente en formato digital.

En el primer grupo nos referimos a aquellos que tienen como finalidad recopilar las imágenes de los mapas de atlas publicados en papel con el objetivo de conservarlos y ofrecerlos a los usuarios en internet tal como se publicaron. Son ejemplos⁴ de este

lingüística porque permite proceder igual que el geólogo lo hace con la corteza terrestre para poder trazar la historia del planeta y estudiar la historia de los usos lingüísticos a partir de los mapas.

³ Algunos investigadores (Bonilla y Bernal 2020: 353) se refieren solo a dos grupos de proyectos de atlas lingüísticos que toman como base las tecnologías de BDE (Bases de datos espaciales) y SIG: «a) atlas lingüísticos diseñados desde su inicio mediante el modelamiento de BDE y el uso de herramientas SIG y b) atlas lingüísticos que en principio fueron análogos y que mediante el procesamiento informático hoy cuentan con una versión digital».

⁴ No se pretende ofrecer una lista exhaustiva de herramientas porque, entre otras razones, resultaría muy difícil (si no imposible) debido a la gran cantidad de recursos que tenemos disponibles en internet actualmente. Los atlas y corpus que se mencionan como muestra sirven para ilustrar los tipos de herramientas que se han desarrollado en el ámbito de la geolingüística digital; y se refieren,

tipo el proyecto de digitalización de los primeros mapas lingüísticos publicados por de Georg Wenker en 1881 y de todos los materiales del *Sprachatlas des Deutschen Reichs* (DSA) —actualmente consultable en RegionalSprache.de; que es el proyecto sucesor del inicial *Digital Wenker-Atlas* (DIWA), cfr. Herrgen 2010; Limper, Pfeiff y Williams 2020: 3744—; el *Sprach und Sachatlas Italiens und der Südschweiz* de Karl Jaberg y Jakob Jud conocido como AIS (accesible desde la herramienta NavigAIS, cfr. Tisato 2019); el *Atlas Lingüístic del Domini Català* (ALDC) y el *Petit Atlas Lingüístic del Domini Català* (PALDC)⁵; el *Atlas lingüístico y etnográfico de Colombia* (que en el marco del proyecto ALEC Digital ofrece la posibilidad de consultar el ALEC facsimilar); y, entre otros, el *Atlas Lingüístico de Puerto Rico* (ALPR), cuyos 75 mapas escaneados pueden consultarse en línea gracias a un proyecto desarrollado por el prof. Enrique Pato financiado por el Ministerio de Educación de Quebec. Mientras algunos de estos proyectos no ofrecen posibilidades de búsqueda (ALDC, PALDC, ALPR) y solo permiten acceder a los datos del contenido de los mapas realizando una consulta como la que suele hacerse al buscar en un atlas en papel —es decir, el usuario examina el atlas por volumen, campo semántico y concepto y, posteriormente, lee e interpreta el mapa—, otros permiten realizar indagaciones específicas a partir de bases de datos espaciales que están conectadas a los mapas digitalizados.

Así, por ejemplo, en el caso del *NavigAIS* es posible buscar por mapa y por punto de encuesta y la aplicación informática recupera el mapa original y ubica el punto de encuesta que se seleccione en el mismo (figura 3).

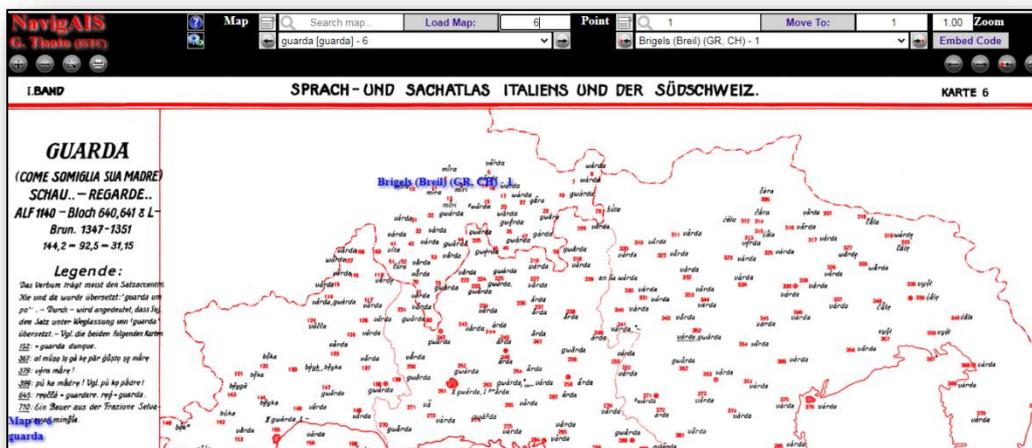


Figura 3. Búsqueda por mapa del *NavigAIS* (ejemplo de un fragmento del mapa 6 - Guarda).

especialmente, como se podrá ir comprobando en los distintos epígrafes de este artículo, a lenguas y dialectos europeos (con mayor representación de las variedades de la península ibérica) y al español de América.

⁵ El primer volumen de este atlas de reciente creación, que se publica tanto en papel como en formato digital, ve la luz en 2007 y se concibe como «un atlas lingüístic de segona generació, que va més enllà de proporcionar mapes puntuals amb la transcripció de les respostes de cada localitat de la xarxa: el PALDC ofereix mapes interpretatius, de tipus areal, que a través d'un codi de colors i sobre un fons poligonitzat permeten fer-se una idea ràpida i acurada de la configuració de la variació dialectal per a cada enunciat i, a més, cadascun va acompanyat d'un comentari que conté l'estudi etimològic, històric i geolingüístic de les diverses variants, elaborat amb sintètica precisió» (Massanell 2017a: 153).

Por otro lado, existen proyectos en los que, además de ofrecerse en internet un mapa con los resultados de los cuestionarios lingüísticos, es posible también hacer consultas a las bases de datos que van asociados al cartografiado digital. Ejemplos de este tipo son el proyecto en el que se está trabajando sobre el *Atlas Lingüístico de la Península Ibérica* ([ALPI](#); García Mouton 2010b; Sousa en prensa) y que actualmente ofrece los resultados «correspondientes a diez preguntas para todos los puntos de la red de encuesta, en transcripción fonética y en ortografía normal. Los resultados de la consulta pueden cartografiarse desde allí» (García Mouton 2017: 340). La base de datos de este atlas de gran dominio puede consultarse de forma conjunta, por localidades, por provincias, por grandes zonas o combinando diferentes posibilidades:

Lo primero que se obtiene son todas las respuestas en transcripción fonética y en la ortografía habitual de cada dominio lingüístico, acompañadas de su localización, su adscripción semántica a los campos establecidos y cualquier anotación de los encuestadores. Se puede acotar la consulta para referirla a un lugar concreto o a varios, a una provincia o a varias, a una zona determinada o a más de una a la vez. (García Mouton 2017: 340)

También disponible en línea —así como en papel y en CD (Aurrekoetxea 2006 [2002])— la versión digital del *Euskararen Herri Hizkeren Atlas* ([EHHA](#)) permite consultar los mapas y las bases de datos asociadas a ellos por campos que ofrecen diferentes tipos de información, como la forma, el ámbito semántico, el número y título del mapa y el punto de encuesta (*zer, gaia, mapa, lekua*). Además, una vez se accede al mapa, es posible ver todas las formas que contiene (*Emaitza mapa honetan*) y junto a cada una de ellas se halla el nombre del punto de encuesta (*Herria*), el lema que le corresponde (lema), la transcripción fonética (*Erantzuna (fon)*) y el audio asociado a ella (*Audioa*). Una de las particularidades que ofrece este atlas es que los mapas que se publican han sido interpretados previamente por los expertos y se presentan divididos en áreas lingüísticas mediante colores; como sucede en el *PALDC*.

Otro proyecto de este tipo es [CartoDialect](#), en el que se ha trabajado sobre la digitalización del *Atlas Linguistique de la France* (*ALF*), y que, además de brindar los mapas originales escaneados desde diferentes tipos de búsqueda (en el apartado *Recherche de cartes* y en el apartado *Visualiseur*), permitir su descarga (cfr. figura 1) así como su búsqueda por nombre y número, también ofrece la posibilidad de consultar las bases de datos asociadas a las informaciones contenidas en ellos. En el apartado *Recherche grammaticale*, el usuario puede buscar por categoría gramatical (*Critère 1*) —Adjectif, Adverbe, Conjonction, Déterminant, Nom, Préposition, Pronom, Verbe— y por las propiedades asociadas a estas categorías (*Propriété*): *Genre (Féminin, Masculin)*, *Groupe*, *Mode (Conditionnel, Impératif, Indicatif, Infinitif, Participe, Subjonctif)*, *Nombre (Pluriel, Singulier)*, *Temps (Futur, Imparfait, Présent, Passé)*. En el tercer tipo de búsqueda, llamada *Recherche thématique*, es posible hacer pesquisas por campos semánticos, algo que supone una actualización de los datos del *ALF*, ya que este atlas se organizó alfabéticamente, a diferencia de la mayoría de sus sucesores, que se concibieron desde una perspectiva onomasiológica a partir del *AIS*. Las búsquedas se dividen en siete grupos semánticos: (1) *Actions, activités des êtres vivants*, (2) *Animaux*, (3) *Caractéristiques physiques*, (4) *Culture*, (5) *Être humain*, (6) *Flore*, (7) *Habitat, vie quotidienne*, (8) *Nature*, (9) *Organisations, pratiques sociales* y (10) *Outils*. Finalmente, en último lugar, el programa permite la búsqueda por punto de encuesta (*Recherche de points d'enquête*). Se recogen datos de los 991 puntos de

encuesta que incluyen información sobre el número identificativo, la localidad, el país (ya que, además de Francia, también se encuestaron pueblos italianos y suizos), el informante (donde se pueden leer las informaciones sobre la edad, la profesión, el estado civil, su vínculo con la localidad que es objeto de encuesta y la fecha de participación en la encuesta) y datos sobre la localidad en un apartado titulado *Liens Externs* que redirige la consulta a la web [DBpedia](#) —que se nutre de las informaciones atesoradas en Wikipedia— donde se sitúa la localidad en el mapa y se pueden leer algunas de sus principales características.

En este mismo grupo se puede situar la página web de los *Índices do Atlas Lingüístico Galego (IALGa)* en la que, aunque no pueden descargarse los mapas, es posible consultar los datos de las formas que contienen los volúmenes publicados. La búsqueda puede realizarse por forma y por el nombre del mapa. En el primer caso, se ofrecen todas las ocurrencias de la forma que se haya buscado y se informa sobre el mapa, el número, el volumen y la página del atlas en la que se encuentran. Adicionalmente, en la parte superior derecha de la pantalla, se proporciona un gráfico de barras en el que puede verse cuál es el volumen en el que se recogen más ejemplos de la palabra que se ha buscado. Así, por ejemplo, en el caso de *man* (figura 4), los Índices recogen 20 formas en 14 mapas, la mayoría de las cuales se hallan en el volumen quinto, que es en el que se sitúan las preguntas relativas al vocabulario del cuerpo humano:



Figura 4. Búsqueda de la forma *man* en los *Índices do Atlas Lingüístico Galego*

Aunque los índices no proporcionan el acceso a la visualización del atlas, algunos de los mapas del *ALGa* son accesibles digitalmente porque se han publicado en la cuenta de [Twitter](#) del *Instituto da Lingua Galega* (ILG). Esto permite a cualquier usuario interesado acceder a algunos de los datos (figura 5) que se van publicando con fines divulgativos:

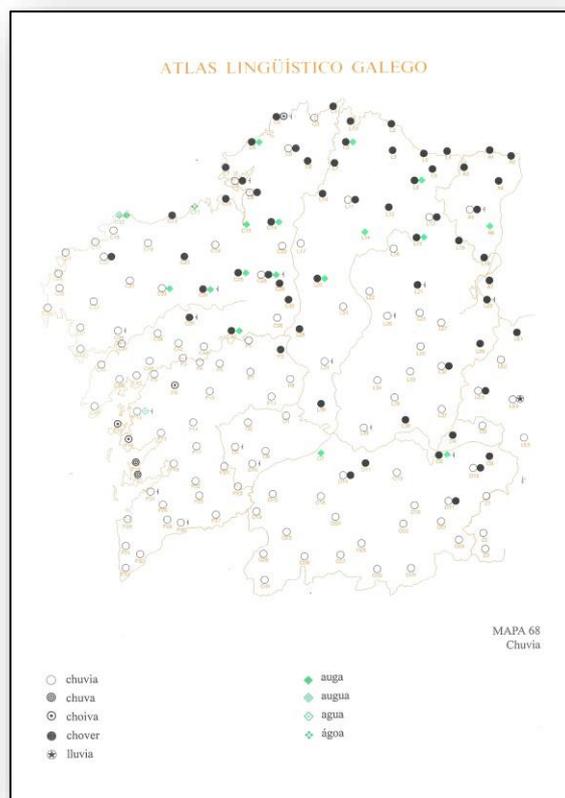


Figura 5. Mapa 68 sobre los nombres del concepto 'chuvia' del ALGa (publicado el 12 de diciembre de 2019 en Twitter)

Sería especialmente útil que la página de los *Índices* permitiera el acceso a los mapas en los que se hallan las formas para poder interpretarlos en relación con el espacio.

Dentro de este grupo se halla también la versión digital del *ALEC* a la que nos hemos referido anteriormente y que, además de ofrecer la consulta de los mapas originales, permite interrogar las bases de datos a las que se han trasladado los contenidos del atlas (Fernández Campos 2017). El usuario puede realizar las consultas desde diferentes puntos de vista: *Consulta especializada*, *Consulta por mapa* o *Consulta por palabra*. En la primera opción, en la búsqueda especializada es posible consultar los datos por uno de los seis tomos y, una vez seleccionado el número de volumen, es posible delimitar el campo semántico. Así, por ejemplo, si se elige el Tomo 1, la aplicación permite acotar la exploración a uno de los tres temas siguientes en los que se distribuye la información: El Campo – Cultivo y Otros vegetales; Industrias relacionadas con la agricultura; o Tiempo y espacio. Una vez seleccionado el tema, se tiene que escoger el mapa, que va identificado con el nombre del concepto al que se refiere. Una vez seleccionado el mapa, se obtiene la lista de variantes léxicas en transcripción ortográfica en la parte derecha de la pantalla, y junto a cada una de ellas se muestra el símbolo que se empleó en el mapa para representarlas. En la parte central de la pantalla, se visualiza un mapa en blanco en el que se pueden incorporar las variantes denominativas según las preferencias del usuario. En la figura 6 se pueden ver representadas, a modo de ejemplo, cinco de las doce variantes léxicas que

se recogieron en el mapa del concepto 'cocotero' (Tomo 1, El campo - cultivo y otros vegetales):

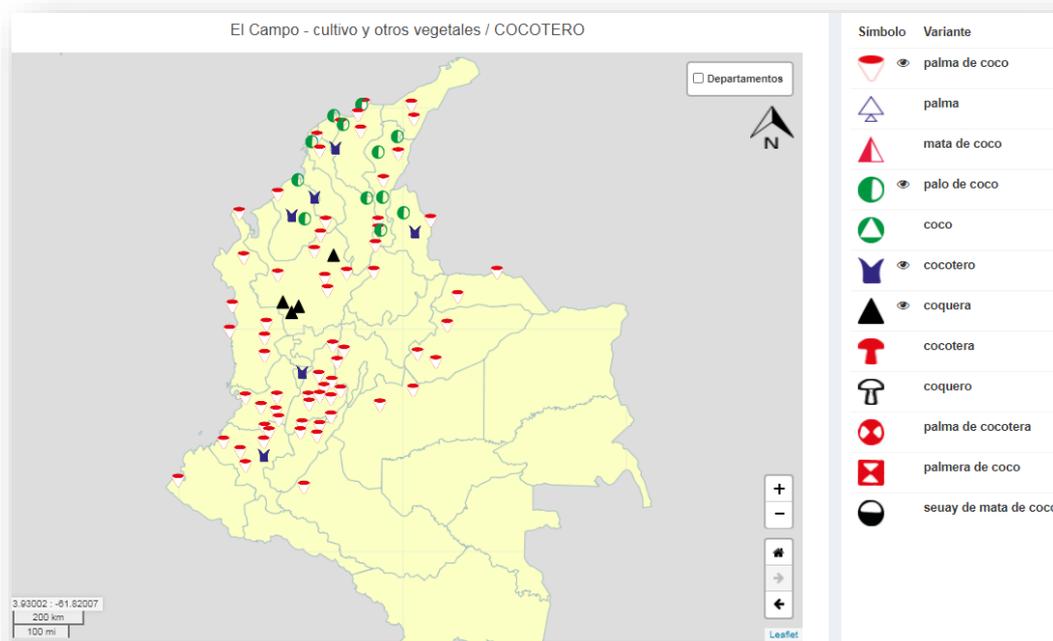


Figura 6. Resultados de la búsqueda *especializada* en el *ALEC Digital* (Tomo 1, El campo - cultivo y otros vegetales, mapa 'cocotero')

Las otras consultas que ofrece el atlas permiten realizar búsquedas por el concepto (*Consulta por mapa*) y por la forma léxica (*Consulta por palabra*).

Las herramientas mencionadas anteriormente son todas consultables a través de internet; sin embargo, según se ha podido comprobar, existen otras similares que también se han cartografiado digitalmente, pero que no pueden consultarse en línea por el momento. Es el caso del *Atlas Lingüístico de Cuba (ALCu)*, cuyos cinco volúmenes se han publicado en DVD (García González 2018). Además, a partir de la edición en soporte digital de estos se han obtenido diferentes aplicaciones informáticas publicadas en la intranet del Instituto de Literatura y Lingüística (ILL) "José Antonio Portuondo Valdor" (Montero Bernal *et al.* 2018): la base de datos geolectal del *ALCu*, el *Atlas Lingüístico de Cuba-Dinámico (ALCu-Din)*, diseñado en ambiente web) y el *Repertorio lexicográfico del Atlas Lingüístico de Cuba* (en construcción). La base de datos permite búsquedas por campos referidos a las distintas informaciones que ofrece el atlas. Sería interesante y productivo poder disponer de estos materiales en línea en un futuro para poder contrastarlos con los de otras zonas del español.

Finalmente, en el proceso de digitalización en el que la geolingüística está inmersa se hallan los proyectos que se han publicado directamente en formato digital y que no han sido previstos para ser publicados en papel. Ejemplos de este tipo son dos atlas regionales de la península ibérica, el *Atlas Lingüístico y etnográfico de Castilla-La Mancha (ALeCMan)* y el *Atlas Dialectal de Madrid (ADiM)*, ambos accesibles en internet. Estos pueden consultarse, igual que los mapas tradicionales, buscando por el campo semántico en el que se encuentra el concepto sobre el que quiere

interrogarse. Su característica principal es que no existe la posibilidad de consultarlos en papel. En este grupo podría incluirse el *Atlas Lingüístico y Etnográfico de Chile (ALECh)*, que se crea a finales del siglo xx con el objetivo de que sea publicado tanto en papel como en versión digital. Sin embargo, actualmente, solo es posible consultar la edición digital (preparada entre 2014-2015), según la información que consta en la web del propio atlas, puesto que para la edición impresa existe solo un prototipo que terminó de prepararse en 2012. Además, la versión digital no puede consultarse de forma completa, pues solo es accesible de forma gratuita una versión demo (figura 7) que ofrece la posibilidad de ver uno o varios mapas para cada una de las secciones. Una parte interesante de esta herramienta digital es que en la parte izquierda de la pantalla se puede leer —a la vez que se observa el mapa— información sobre el punto de encuesta, los informantes, la fecha en la que se realizaron las encuestas, los encuestadores que realizaron el trabajo de campo y en qué otros atlas del español se puede encontrar el concepto (como es habitual en los atlas impresos).



Figura 7. Imagen sobre el mapa 11 'garganta' de la versión demo del ALECh

Según García Mouton (2016a: 258), el ALECh es un atlas amigable porque no está pensado para ser consultado solo por especialistas, ya que se combinan las respuestas fonéticas con su transcripción ortográfica y con símbolos, igual que en el ALEC.

Uno de los inconvenientes de este tipo de atlas (*ALeCMan*, *ADiM*, *ALECh*), si se contrastan con algunos de los anteriores (*ALEC Digital* o *CartoDialect*), es que no pueden lanzarse consultas sobre sus datos, por lo tanto, la explotación sigue basándose en la lectura e interpretación que el interesado realice sobre los mapas. A pesar de ello, de acuerdo con García Mouton (2016a: 257), cualquier edición digital sigue teniendo enormes ventajas frente a la clásica impresa: «las ventajas de un atlas informático en línea son evidentes, ya que lo hacen accesible desde cualquier computadora desde los lugares más alejados de las bibliotecas especializadas y, además, las posibilidades de una edición exhaustiva, sin límites de espacio, de toda la información recogida en las encuestas, son mucho mayores».

2.3. Un nuevo concepto de atlas lingüístico: ¿atlas o geocorpus?

En la actualidad, el investigador dispone de muchas y muy variadas herramientas digitales de consulta en línea para el estudio de la variación lingüística. [Dialectoteca del español](#) (University of Iowa) o la aplicación [Dialectos del español](#) son una muestra de este tipo de recursos. El primero se define como una biblioteca audiovisual que documenta los patrones más representativos de la pronunciación del español moderno; y el segundo, de reciente creación, es una web-app sin ánimo de lucro destinada a conocer la variación de la lengua española en la actualidad (Bouzouita, Castillo y Pato 2018) y basada en app similares creadas para los dialectos suizo-alemanes ([Dialäkt Äpp](#) y [Voice Äpp](#)) y para los dialectos del inglés de Inglaterra ([English Dialects App](#)). En muchos otros casos, también con el objetivo del estudio de la variación dialectal, se han creado corpus de datos que utilizan los mapas digitales —aunque de formas diversas y con objetivos variados— para situar geográficamente las informaciones lingüísticas; el *Corpus Oral y Sonoro del Español Rural* ([COSER](#)), el *Corpus Dialectal del Extremeño* ([CoDiEx](#)), el corpus oral del proyecto *Frontera Hispano-Portuguesa* ([FRONTESPO](#)), el *Tesouro do léxico patrimonial galego e portugués*, la web [Estratigrafía Dialectal](#) del catalán o la base de datos terminológica y de identificación de especies pesqueras de las costas de Andalucía ([Ictioterm](#)), son solo una muestra del gran abanico de aplicaciones que compilan datos lingüísticos y los sitúan en el eje diatópico.

Algunas de las aplicaciones, como se ha indicado anteriormente, emplean el término *atlas* en su denominación a pesar de que se alejan casi por completo del concepto tradicional de este tipo de compilaciones desde una perspectiva metodológica (Julià 2020). Los datos que se representan y recogen en este tipo de plataformas proceden de fuentes variadas como encuestas —en las que no se aplica el método de la geolingüística y, en muchas ocasiones, se realizan por internet—, entrevista, vaciados de obras lingüísticas o de textos extraídos de redes sociales como Twitter. Además de contar con datos de distinta procedencia, muchas de las herramientas tienen un objetivo más específico que el de los atlas tradicionales, pues se centran en un aspecto lingüístico. Son una muestra de este tipo los proyectos que se describen a continuación (vid. nota 4).

El creciente interés por el estudio de la variación sintáctica en la lingüística del siglo XXI se ha evidenciado no solo en el aumento de las investigaciones dedicadas a este asunto, sino también en el número de corpus que tienen por objeto de investigación este tema. Entre ellos, queremos destacar los que han elegido la cartografía digital como medio principal de representación de los datos y que, por ello, se han etiquetado como atlas lingüísticos. Existen ejemplos para el escocés (*Scots Syntax Atlas*, [SCOSYA](#)), el español (*Atlas Sintáctico del Español*, [ASinEs](#)), el italiano (*Atlante Sintattico d'Italia*, [ASIt](#)) o el neerlandés (*Dynamic Syntactic Atlas of the Dutch dialects*, [DynaSAND](#)), entre otras lenguas. El *ASinEs* (*Atlas Sintáctico del Español*), que es una aplicación accesible en línea con formato de atlas dedicada al estudio sincrónico de la variación sintáctica de los geolectos del español europeo y americano (Cerrudo *et al.* 2015), se presenta como un trabajo pionero basado principalmente en la recopilación de los datos de dos gramáticas de referencia (*GDLE* 1999 y *NGLE* 2009); aunque se ha previsto completar los datos de estas obras con otros procedentes de encuestas y de las redes sociales. El atlas surge como proyecto de investigación

coordinado por Ángel Gallego desde la Universitat Autònoma de Barcelona y en convenio con la Real Academia Española.

En el área de los estudios fonético-fonológicos son también muy interesantes los atlas consultables en línea que se han publicado sobre las variedades románicas. El [Atlas interactivo de la entonación del español](#) tiene como objetivo presentar materiales para el estudio de la prosodia y la entonación de las variedades del español a través de su variación en el espacio para poder realizar estudios comparativos. Los materiales que contienen se hallan, a diferencia de los de muchos otros atlas de nueva generación, en formato de audio y vídeo. En esta línea se hallan los distintos atlas publicados para el macroatlas tománico titulado *Atlas Multimedia de la Prosodia del Espacio Románico* ([AMPER](#)) consultables en el espacio web [Geoprosodic](#) e impulsados desde la Université Stendhal Grenoble III por Michel Contini y Antonio Romano. Estos se centran en la recopilación de datos de ocho dominios lingüísticos (español, catalán, gallego, asturiano, galorrománico, italarrománico, el portugués y el rumano). Se trata de una investigación internacional cuyo objetivo principal es el estudio de la entonación y el acento. En la página Geoprosodic se destaca el enfoque con el que está concebido como una de las propiedades de este gran proyecto:

principalmente fonético, con implicaciones fonológicas, dialectológicas y sociolingüísticas. Lo que significa que pertenece a esta nueva era de la geolingüística en la que no sólo se amplía el campo de estudio, en este caso, al nivel suprasegmental, sino que además se tienen en cuenta todas las variables (ámbito geográfico, sexo, edad, nivel de instrucción, procedencia, etnia, etc.).

En el ámbito de los estudios léxicos se sitúan algunas de las primeras herramientas en las que se ofrecían mapas lingüísticos digitales en la red. Destacan, por ejemplo, los resultados obtenidos en el desarrollo del proyecto *Variación léxica del español en el mundo* ([VARILEX](#)), que se inicia en la última década del siglo xx, coordinado por el profesor Hiroto Ueda (Universidad de Tokio), con el objetivo de recopilar, principalmente mediante encuesta, datos sobre el vocabulario del español urbano en el mundo. Aunque el método de encuesta se alejaba del empleado en la geografía lingüística, permitía visualizar los resultados en un apartado de la página del proyecto titulado *PHP-Atlas Varilex* (figura 7). Inicialmente, la aplicación daba la posibilidad al usuario crear mapas de las informaciones que las bases de datos almacenaban mediante cartografía automática (Chacón 2016); lamentablemente, hoy esta opción —que era la que acercaba los resultados de la investigación a la geolingüística, como puede advertirse en el mapa de la figura 8— ha desaparecido en la nueva web del proyecto.



Figura 8. Mapa de *Varilex* (consultado en 2013) sobre la designación *anteojos* para el concepto 'gafas'

La posibilidad de descargar este tipo de mapas, que informaban sobre la presencia y la ausencia en el mundo hispanohablante de una determinada forma léxica, ha sido sustituida por una aplicación web creada en el marco del desarrollo de una tesis doctoral en la que se tomaban los datos de *Varilex* como corpus de estudio (Chacón 2016). En esta aplicación, llamada *Grado de representatividad léxica del español (GRP)*, se abren los mapas en los que se documentan las variantes léxicas (figura 9):



Figura 9. Mapa sobre las designaciones del concepto 'blackboard' en español obtenidas en Varilex y representadas en la web GPR (Chacón 2016)

La sustitución de los mapas de la figura 8 por los que se ofrecen en la web GPR (figura 9) supone una mejora tanto en la visualización de los datos (que se ofrecen en bruto en la parte inferior del mapa) como en el valor que se les da en su representación.

Con un enfoque más general y alejado del método de la geolingüística, se pueden encontrar propuestas innovadoras como la del *Atlas Dialectal de Barcelona (ADiBa)*, un proyecto de reciente creación que se define como una herramienta —un microatlas centrado en el Área Metropolitana de Barcelona— para explorar la variación lingüística del español de Barcelona y que parte de la información que ofrece la red (Twitter), de la colaboración ciudadana (*crowdsourcing*) y de entrevistas (Arias y Gallego 2019). Esto es lo que aleja este atlas del antiguo método geolingüístico y lo integra en el marco de las más recientes investigaciones en variación lingüística que toman como corpus de estudio datos procedentes de internet. Esta interesante perspectiva metodológica está generando múltiples proyectos tanto por la velocidad a la que se obtienen los datos como por el volumen de los mismos (Ruiz Tinoco 2018: 599). Sobre el uso de los macrodatos (*big data*) procedentes de la red para el estudio de la variación lingüística, Estrada y De Benito (2016b: 8) ofrecen una serie de reflexiones muy interesantes sobre su utilidad, sus posibilidades y «las precauciones que debe tomar el lingüista al enfrentarse a ellos». Este tipo de trabajos se hallan en lo que se ha denominado dialectología aglomerativa (*Aggregate Dialectology*), ya que, según Aurrekoetxea (2019: 22), se enmarcan en «una dialectología cimentada en el valor del uso de grandes masas de datos, en vez de determinadas características escogidas subjetivamente por el dialectólogo», lo que, según el mismo investigador:

rompe claramente con el modo tradicional de investigación para entrar en una práctica investigadora nueva y desconocida, gracias al empleo de métodos estadísticos. La dialectología aglomerativa se fundamenta, pues, en dos pilares: el uso de grandes masas de datos y la cuantificación automática para la medición de las diferencias lingüísticas entre localidades y la clasificación jerárquica de las hablas (Aurrekoetxea 2019: 22).

Aunque los datos procedentes de la web pueden ser comparables a los que se puedan recoger mediante el método de encuesta habitual en la geografía lingüística, debe tenerse en cuenta que estamos ante un tipo de lenguaje diferente que procede de una nueva forma de interactuar y de comunicarse, la denominada comunicación mediada por ordenador o computadora (Estrada y De Benito 2016a; 2016b; 2018). Según Ruiz Tinoco (2018: 599), la lengua empleada en las redes sociales es una lengua distinta porque no es ni lengua oral ni lengua escrita. Además de esta diferencia, que es esencial, existen otras cuestiones que hay que tener en cuenta si se comparan los atlas tradicionales con los atlas (o geocorpus) del siglo XXI en los que se toman los datos de la web. En primer lugar, en muchos casos en los datos de la red se desconocen las informaciones personales (edad, sexo, lugar de nacimiento, lugares de residencia, información lingüística de los padres y otros familiares) y académicas de los hablantes (formación recibida), algo que puede condicionar las producciones escritas que puedan dejar en las redes sociales. En segundo lugar, aunque puedan rastrearse algunos de los datos de los perfiles de los que se hayan extraído, la veracidad de los mismos es difícilmente contrastable; por ello, puede resultar complejo determinar los perfiles lingüísticos que se están comparando. En tercer lugar, no es posible acotar los temas sobre los que se pueden encontrar informaciones y se limitan a lo que los usuarios teclean en las redes sociales; por lo tanto, el investigador no obtiene datos relativos a aquello que le interesa, sino a aquello que le ofrece la red. Esto se opone al método tradicional en el que la planificación de la encuesta es una de las partes más importantes, pues gracias a ella se puede llevar a cabo el análisis previsto de antemano (García Mouton 2016b: 34-35). A pesar de las diferencias entre los datos, los informantes y el método de obtención de los mismos, los dos tipos de atlas (tradicionales y modernos) comparten algunos rasgos. El texto que se vuelca en una conversación o un mensaje de Twitter o Instagram puede contener errores de muy diverso tipo (ortográficos, de morfosintaxis, de cohesión, de puntuación, etc.), involuntarios o no, para mostrar un tipo de lengua que puede no reflejar el comportamiento lingüístico habitual del hablante. Esta cuestión, entre otras, ha sido también criticada en la geolingüística tradicional (Caravedo 2005).

Por ello, siguiendo a Ruiz Tinoco (2018), en la investigación sobre variación lingüística parece totalmente necesario mantener el método de encuesta tradicional; es decir, no puede pretenderse que la recopilación automatizada de datos procedentes de las redes sociales sustituya a la metodología convencional, en cualquier caso, puede ser un complemento:

Es necesario aclarar que la recogida de datos para el estudio de la variación lingüística se puede realizar de varias maneras y que la recogida a través de Internet, ya sea de blogs, prensa o redes sociales, no pueden [sic.] sustituir a la investigación tradicional por medio de encuestas, ya que éstas tienen muchas ventajas gracias al contacto directo del investigador con los informantes, cuyos atributos se pueden determinar de antemano según el objetivo de la investigación (Ruiz Tinoco 2018: 598).

Un proyecto en el que se aplica la metodología de la geolingüística tradicional adaptada a la era digital y basada en la recopilación de datos a través de internet es [Tweetolectology](#) (University of Cambridge), un perfil de la red social Twitter desde el que se publican mapas lingüísticos del Reino Unido. Aunque en su denominación no se identifica como un atlas, es interesante tener en cuenta que se trata de una base de

datos lingüísticos que ofrece la información representada en mapas. Los mapas son de dos tipos: por un lado, el usuario puede encontrar cartas lingüísticas que representan datos que proceden de encuestas actuales realizadas en línea y, por otro, mapas publicados a mediados del siglo xx en el marco de la geografía lingüística. Se trata, por tanto, de una herramienta en la que se pretende ofrecer una visión histórica de determinados fenómenos lingüísticos y servir para el estudio del cambio lingüístico. Los datos que se publican son de diversa procedencia y pertenecen a fuentes distintas que permiten estudiar la evolución lingüística desde diferentes perspectivas:

Most maps we post on this account are from four sources- the Survey of English Dialects (SED), a 1950s survey of rural speech in England; the English Dialects App (EDA), a 2016 smartphone-based survey of speech in the British usage on Twitter itself, collected as part of our project; and informal surveys of this account's followers which we'll be posting here. Look at the top-right corner of each map to see the source (*Tweetolology*, 20 de julio de 2019).

A continuación, se muestran dos mapas lingüísticos sobre el mismo concepto que se corresponden con épocas distintas y cuyos datos proceden de fuentes diferentes. El mapa de la figura 10 recoge los nombres informales para referirse al concepto 'abuelo' que se recopilaron a mediados de siglo xx en la *Survey of English Dialects* (SED) dirigida por el Harold Orton (University of Leeds) con el objetivo de la geolingüística original (Klemola y Jones 1999), esto es, documentar los datos sobre los dialectos de Inglaterra antes de que se perdieran:

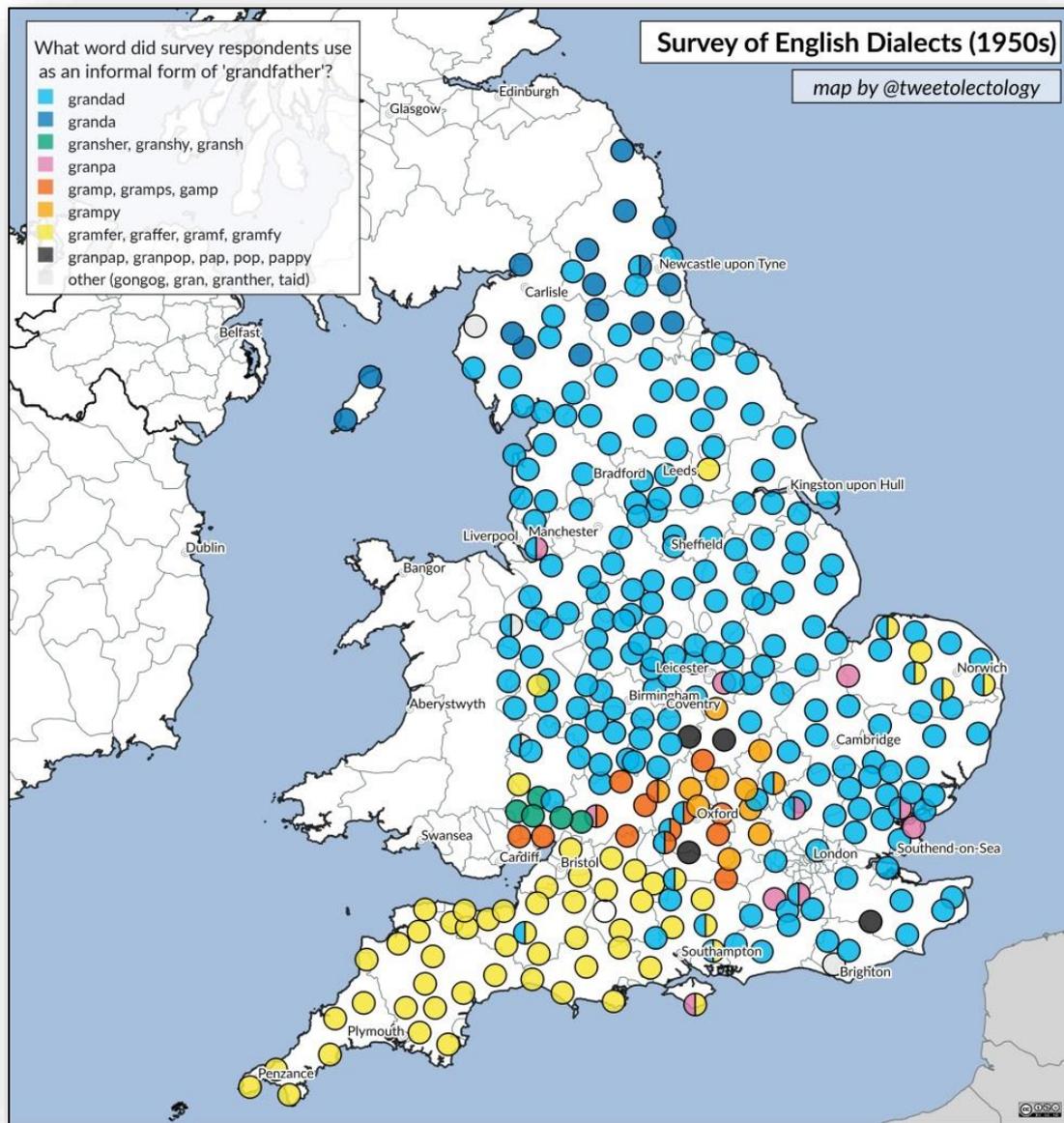


Figura 10. Mapa léxico sobre las designaciones coloquiales del abuelo publicado en @Tweetolectology y procedente de SED (1950)

Estos datos se pueden contrastar con los que se ofrecen a partir de la encuesta que se ha realizado específicamente desde la [survey platform for @tweetolectology](#) para el mismo concepto y que permiten generar mapas como el de la figura 11.

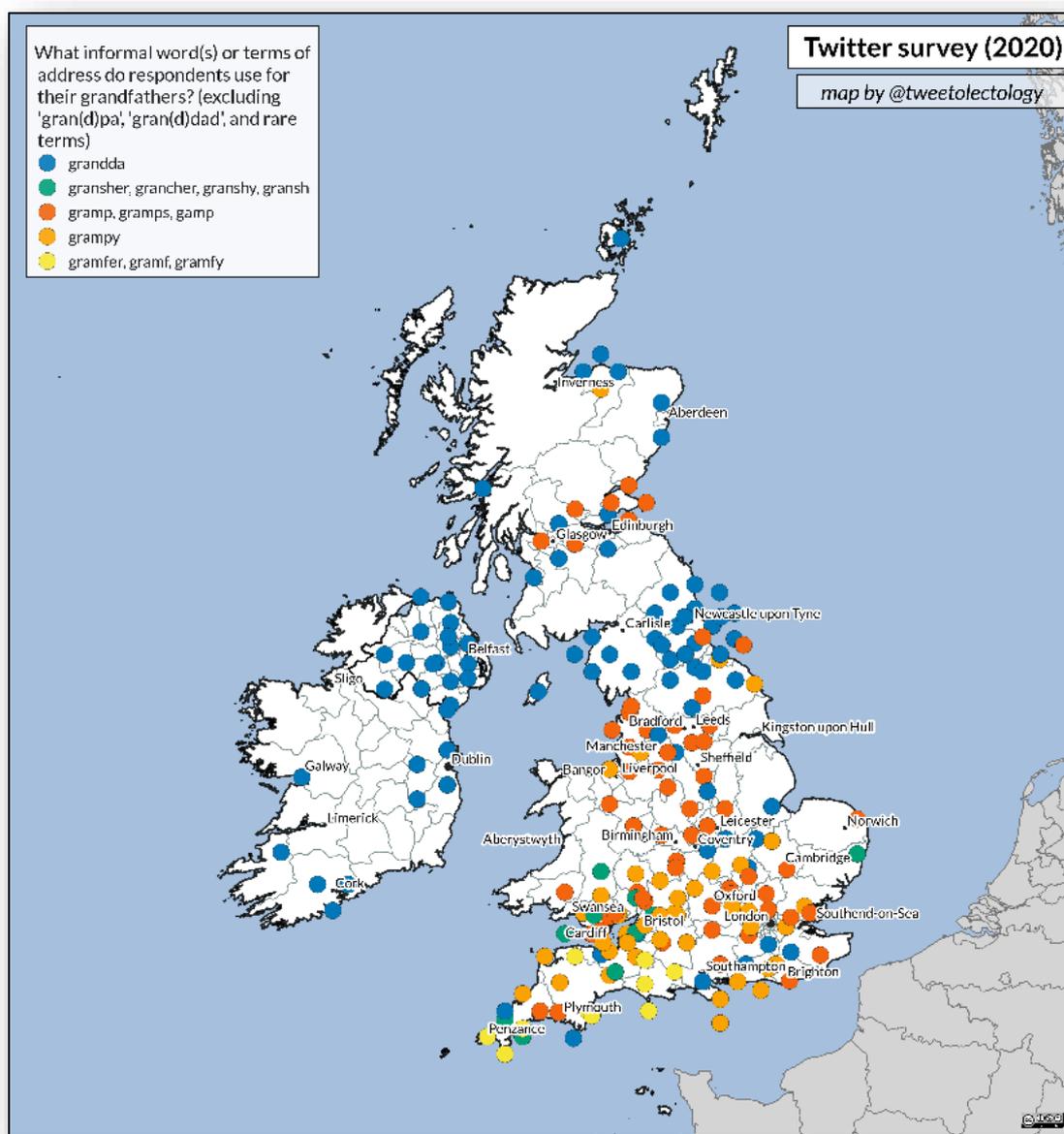


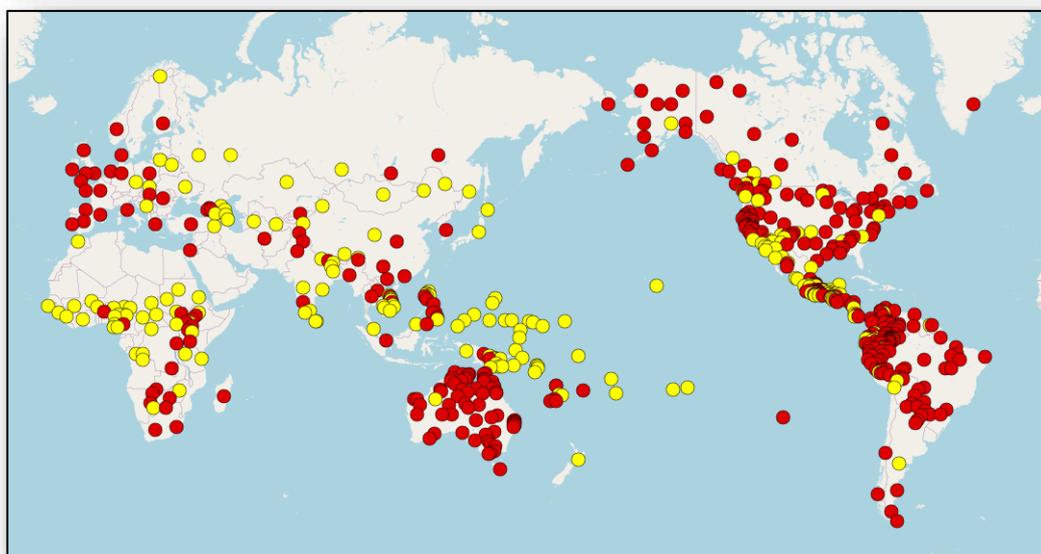
Figura 11. Mapa léxico sobre las designaciones coloquiales de los abuelos procedente de la encuesta en línea y publicado en @Tweetolectology

Aunque las preguntas que se pueden encontrar representadas en los mapas de esta herramienta se refieren a aspectos lingüísticos de diversa índole (fonética, morfología, sintaxis y léxico), el proyecto de mapear los datos de esta red social nace con el objetivo de estudiar la variación morfosintáctica y la extensión de los cambios en el tiempo:

We are interested in the way in which changes in the grammar of languages ('innovations') spread out from a small number of speakers to a larger section of the population ('diffusion'). We use Twitter to gather large quantities of localisable data from many places and across large areas; this allows us to investigate language variation and change in fine geographical detail (About the Project, [Tweetolectology](#))

Como se deriva del contraste de las figuras 9 y 10, uno de los aspectos más interesantes de este proyecto es el contraste de procedencias de los datos, pues se compilan informaciones que pueden mantener una distancia de hasta 100 años entre ellas. Las nuevas tecnologías facilitan el estudio de la evolución de la lengua y los cambios lingüísticos; permiten también comprobar cuál es su distribución en el espacio y difundir de forma rápida y sencilla los resultados entre los propios agentes del cambio, los hablantes.

En esta línea es muy interesante, por tratarse de una herramienta que proporciona visión genérica sobre las lenguas del mundo, el *The World Atlas of Language Structures online (WALS)*. Este se define como una gran base de datos que contiene información lingüística variada (fonológica, morfológica, sintáctica, léxica) de todo el mundo con el objetivo de ofrecer una panorámica de las relaciones que existen entre las lenguas. Cada uno de los mapas va acompañado de un importante número de datos que proceden de las investigaciones de los colaboradores del proyecto. Así, por ejemplo, si se realiza una búsqueda por capítulo (aparecen 152 ocurrencias) y se restringe por ámbito léxico en el campo Área, aparecen 13 entradas, la primera de las cuales se refiere a una cuestión relativa a la categorización del cuerpo humano. El mapa de la figura 12 muestra una clara división entre las lenguas que categorizan el brazo y la mano como una sola unidad —esto es, que emplean la misma palabra para referirse a estas dos partes del cuerpo— y las que lo hacen como dos unidades distintas:



| Value | Representation |
|--|----------------|
| ● Identity: a single word denotes both 'hand' and 'arm' | 228 |
| ● Differentiation: one word denotes 'hand' and another, different word denotes 'arm' | 389 |
| Total: | 617 |

Figura 12. Mapa extraído de WALS (Brown 2013)

Finalmente, no deben dejarse de lado los proyectos de carácter histórico como el *Atlas Lingüístico diacrónico e interactivo de la Comunidad de Madrid* ([ALDICAM-CM](#)) dirigido por Pedro Sánchez-Prieto desde la Universidad Alcalá de Henares que tiene como fin elaborar un mapa diacrónico dinámico de las particularidades del habla de la Comunidad de Madrid. Se trata de un corpus digital histórico que parte de la transcripción de documentos fechados entre el siglo XIII y el XIX procedentes de diferentes fondos documentales que permite proyectar y representar los datos en un mapa. La representación cartográfica digital es un complemento al corpus que ofrece la transcripción paleográfica, una presentación crítica y la imagen del documento transcrito. Igual que en el *ALPI*, el atlas ofrece la posibilidad de que el usuario pueda seleccionar los parámetros de consulta en los que tenga interés y proyectar un mapa a la carta. En esta línea de trabajo se halla también el *Atlas Lingüístico Diacrónico y Dinámico del español* ([ALDIDI](#)), que constituye una de las funcionalidades en vías de desarrollo para el *Corpus de documentos Españoles Anteriores a 1800* (CODEA).

Los datos hasta aquí expuestos constituyen una muestra de los beneficios que aporta la digitalización, el uso de los SIG y BDE al estudio de la variación lingüística en general. Y, en particular, también se ha justificado que la actualización de los atlas lingüísticos tradicionales va acompañada de la creación de un importante número de nuevas herramientas de carácter geolectal.

3. UN CORPUS DE LOS ATLAS LINGÜÍSTICOS

El número de herramientas descritas en el apartado anterior permite afirmar que las posibilidades de las que dispone el investigador para estudiar los dialectos en la actualidad son muy variadas; y todo parece indicar que muy probablemente siga aumentando el número de recursos de este tipo. Es evidente que la digitalización ha generado un creciente interés por la creación de corpus y otras herramientas que ofrezcan la posibilidad de estudiar la variación según su extensión en el espacio. Se ha podido comprobar que, además de los corpus, la última década ha supuesto una considerable mejora en los estudios de la geografía lingüística, pues muchos de los atlas, gracias a las posibilidades que ofrecen los SIG, se han publicado en línea y permiten hoy acceder a los materiales que atesoran sus mapas.

Para el español europeo, variedad que no ha visto nunca publicados por completo los datos de su atlas lingüístico nacional (*ALPI*) —aunque se está trabajando en ello actualmente (García Mouton 2017) y durante un tiempo estuvieron disponibles en internet los materiales de los cuadernos de respuestas que David Heap recopiló y escaneó entre 1999-2001 (Heap 2002)—, parece que la publicación digital de los datos de los atlas regionales supondría un complemento esencial en cualquier estudio de dialectología sincrónica y diacrónica. Los atlas lingüísticos regionales son materiales de incalculable valor e interés cultural, lingüístico y antropológico por cuanto la información que contienen no ha sido analizada de forma completa y por cuanto recogen en sus mapas una parte del patrimonio cultural, lingüístico y cognitivo de la España de la segunda mitad del siglo XX (Del Barrio 2018; Fernández Morell en prensa). Actualmente, el hecho de que estos materiales solo puedan consultarse en formato papel en algunas pocas bibliotecas dificulta sus posibilidades de acceso, divulgación y estudio (Sousa 2017: 18). Este es uno de los grandes problemas que la geolingüística

no ha logrado todavía solventar y del que muchos investigadores se han lamentado (Alvar 1984: 112-113; Aurrekoetxea 2006 [2002]: 63). Mediante el traspaso de las formas lingüísticas de los mapas a bases de datos se favorecería su difusión y aumentarían sus posibilidades de estudio, además de permitir que los materiales permanecieran almacenados con el fin último de contribuir a su preservación. Consideramos que tanto los atlas como su contenido forman parte de los bienes y la historia de la lengua española y que es necesario y urgente conservarlos para evitar que se pierdan. Así, uno de los objetivos con el que se concibieron a finales del siglo XIX —el de ser tesoros que almacenaran datos lingüísticos que se preveía que iban a desaparecer (Navarro Carrasco 1993: 309-310); además de explicar la vida de las palabras y la variación en las hablas vivas (García Mouton 2016b: 31)— debe ser el mismo que impulse su digitalización y su publicación en línea, pues el riesgo de perder de nuevo estos materiales es muy elevado.

Las razones que han motivado la idea de crear un corpus en el que se recojan los datos de los atlas lingüísticos regionales del español son de diverso tipo, aunque pueden resumirse de forma general en tres puntos que responden a las necesidades del investigador del siglo XXI. En primer lugar, se parte de la necesidad de examinar todos los materiales que la geolingüística recopiló y publicó a lo largo de la segunda mitad del siglo XX con mucho esfuerzo y con una gran inversión económica, ya que han sido parcialmente explotados y no se les ha podido extraer el rendimiento para el fueron ideados (García Mouton 2009; Sousa 2017; Julià 2018, en prensa). En segundo lugar, la imposibilidad de acceder a los mapas de forma sencilla y rápida. En este sentido, la informática pone a disposición de los investigadores, en el contexto de las humanidades digitales, una serie de herramientas que permiten almacenar, estructurar y publicar los contenidos de los mapas para que puedan ser estudiados de forma completa. Solo de esta manera será posible extraerles el máximo rendimiento y entender los datos que contienen. En tercer lugar, los datos de los mapas son complementarios a los que ofrecen los corpus textuales y lexicográficos del español que actualmente se pueden consultar en línea y, por lo tanto, constituirán una fuente de información adicional a los estudios que parten de estas herramientas (Fernández Sevilla 1974, Salvador 1980, Alvar 1982, Navarro Carrasco 1990, 1992, 1996, 2001, Saralegui 2005, Prat Sabater 2006, Castañer Martín 2009, García Mouton 1990, 2010a, 2016c; Del Barrio 2018).

En esencia, disponer de los datos de los atlas tradicionales en formato digital abrirá nuevas posibilidades de estudio a la investigación diatópica y diacrónica del español europeo y la pondrá al nivel de otras variedades como el francés (*CartoDialect*), el gallego (*Índices del ALGa*), el euskera (*EHHA*) o el español de Colombia (*ALEC Digital*) que disponen de algunas herramientas que permiten consultar la información lingüística de los mapas desde diferentes perspectivas (§ 2.2).

3.1. Objetivo y metodología

Es en este marco en el que nace la idea de la creación del *Corpus de los Atlas Lingüísticos (CORPAT)*, una herramienta informática cuyo objetivo es la sistematización y organización de las informaciones contenidas en los atlas lingüísticos regionales de la península ibérica (*ALCyL*, *ALEA*, *ALEANR*, *ALECan*,

ALeCMan, *ALEICan*, *ADiM*)⁶. El corpus no pretende convertirse en la versión digital de los atlas regionales, sino constituir un banco de datos que permita acceder a las informaciones de sus mapas en conjunto. Su comparabilidad, más que demostrada en numerosas investigaciones, deriva esencialmente de haberse creado con una metodología común (Del Barrio 2018: 64). Por lo tanto, será un complemento a una posible futura digitalización facsimilar de los mapas.

De la consecución de los objetivos principales se desprenden un conjunto de objetivos secundarios que permitirán contribuir al conocimiento lingüístico del español tanto desde una perspectiva especializada y académica como desde un punto de vista más genérico: (a) divulgar los resultados de la investigación en materia filológica para dar cuenta de qué tipo de investigaciones se realizan en humanidades y cuál es su aplicación real; (b) dar a conocer qué es la geografía lingüística y actualizarla; (c) difundir los estudios en materia geolingüística española de la segunda mitad del siglo xx; (d) facilitar el acceso de la sociedad a la información sobre la variación lingüística; (e) favorecer la difusión de la riqueza de la variedad lingüística del español entre los sectores no especializados (ámbito educativo y turístico); (f) contribuir al estudio de las relaciones léxico-semánticas del español (onomasiológicas-semasiológicas); (g) proporcionar información para el estudio de los mecanismos de creación léxica (metáfora, metonimia); (h) aportar datos para el examen de los patrones de categorización de las lenguas (universales lingüísticos); (i) facilitar el estudio de la variación desde diferentes perspectivas teóricas (por ejemplo, en el ámbito de los estudios léxicos, se puede ver la relación entre la variación y la teoría de prototipos en el marco de la semántica cognitiva); (j) complementar los datos de los diccionarios y los corpus textuales; (k) observar la pervivencia de las formas vernáculas y la relación que se pueda establecer con la forma estándar; (l) analizar los datos desde una perspectiva cuantitativa (dialectométrica) con el fin de poder comprender de forma global la variabilidad de las informaciones que contienen (Goebel 2010).

Tomando como punto de partida estos objetivos, se inició el diseño de *CORPAT* gracias a la concesión de una modesta ayuda económica en la Convocatoria de Ayudas Complementarias a la investigación para el profesorado de la Facultad de Filología (Convocatoria 2019, BICI n.º 13 (14/01/2019), epígrafe 11) para cubrir gastos derivados de la prestación de servicios complementarios a la investigación (corrección ortotipográfica y de estilo, procesamiento de datos, diseño de herramientas informáticas, asesoramiento lingüístico, etc.) (MODALIDAD D). La financiación recibida permitió iniciar el trabajo de diseño del corpus (interfaz de introducción de datos e interfaz de consulta) y de la página web desde la que se difundirán los resultados, que se halla en una fase de desarrollo primario y que seguirá perfeccionándose en los próximos meses.

3.2. Diseño y estructura

Actualmente, gracias a la financiación recibida, se ha diseñado una página web que incluye tanto la información sobre el proyecto como una primera muestra del contenido del corpus.

⁶ Aunque este es el objetivo principal, no se descarta ampliar el número de atlas (tanto sobre el español de América como de otras variedades de la península ibérica) en el futuro.

El corpus se divide de dos interfaces, la de introducción de datos y la de consulta. La interfaz de introducción de datos (administración) permite incorporar la información en línea para que diferentes usuarios puedan trabajar en la implementación de forma simultánea y desde diferentes lugares. Esta base de datos relacional se compone de diferentes tablas que contienen todas las informaciones que se desprenden de la lectura e interpretación de las respuestas recogidas en cada punto de encuesta para cada mapa (*forma, concepto, campo semántico, subcampo semántico, punto de encuesta, localidad, provincia, atlas, lengua, categoría gramatical, información morfológica, información fonética, información semántica, información etnolingüística, número de respuesta*). En estos campos se pretende recoger la mayor parte de la información que incluyen los mapas lingüísticos con el objetivo de ofrecer al usuario la posibilidad consultarla desde diferentes puntos de vista. Se combinan, por ello, ciertas funcionalidades que ofrecen algunas de las herramientas anteriormente descritas, como el *ALEC Digital*, *CartoDialect* o los Índices del *ALGa* (§ 2.2).

El proceso de vaciado de los mapas supone acercarse a ellos desde múltiples perspectivas que permitan al usuario ir más allá de lo que puede apreciarse en las cartas lingüísticas impresas. Por un lado, se vuelca en la base de datos la información que el mapa proporciona: la forma de cada punto transliterada ortográficamente; el nombre del mapa, que se refiere al concepto por el que se preguntó durante la encuesta; el campo y el subcampo⁷ semántico al que pertenece (según la organización onomasiológica del atlas); el punto de encuesta; el tipo de respuesta (si se recoge como primera, segunda o tercera respuesta) en el caso de que existan múltiples respuestas; el número de mapa; el volumen en el que se publicó; el atlas en el que se encuentra; y la información adicional que se haya recogido sobre la respuesta que el informante ha ofrecido (generalmente situada en el margen del mapa). Y, por el otro, los datos derivados de un análisis lingüístico de las variantes recogidas (p. ej. la categoría gramatical de la forma o la información morfológica asociada a ella). Además, con el fin de poder aprovechar las posibilidades que ofrecen los SIG (mediante el programa libre de OpenStreetMap), se introduce la ubicación del punto de encuesta en el espacio.

⁷ Desde el *AIS*, los índices de los atlas lingüísticos regionales suelen clasificar los mapas en campos generales que se dividen en subsecciones semánticas. Así, por ejemplo, en el volumen II del *ALEA*, el apartado primero se dedica al ámbito de los vegetales, el cual, a su vez, categoriza los mapas en subámbitos de especialización: a) plantas silvestres, flores, arbustos; b) hortalizas; c) árboles frutales; d) el bosque.

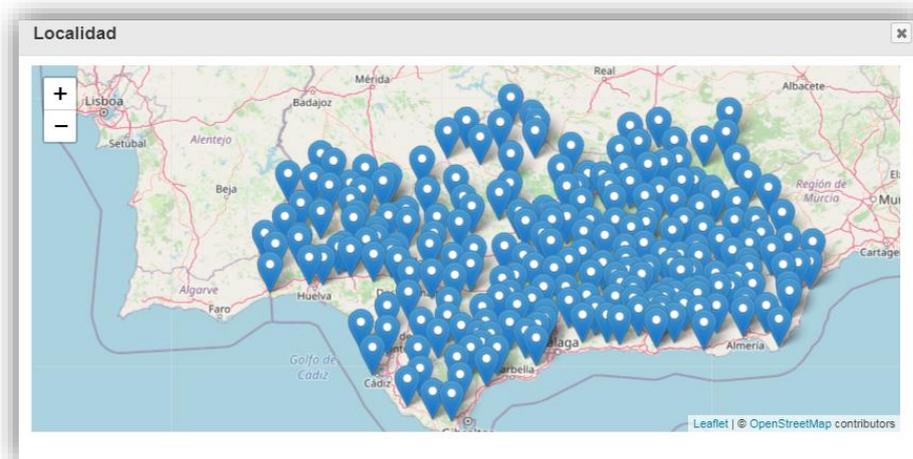


Figura 13. Puntos de encuesta del ALEA desde la visualización de OpenStreetMap

La interfaz de consulta, que actualmente se halla en fase de desarrollo, se podrá interrogar desde el apartado *Consulta* de la página web y permitirá llevar a cabo diferentes tipos de indagaciones según el interés del usuario. La búsqueda abierta arrojará resultados sobre cualquier información sin necesidad de especificar por campo. La búsqueda simple admitirá acotar por los tres campos principales (*forma*, *concepto* y *campo semántico*). Las posibilidades de este tipo de interrogación suponen poder obtener, entre otras opciones, una lista de todos los mapas en los que se encuentra una forma léxica determinada (igual que sucede en la aplicación de los *Índices léxicos del ALGa*) o saber si una misma forma se emplea como designación de realidades pertenecientes a campos semánticos distintos, esto es, acercarse a los datos de los atlas desde la perspectiva de la polisemia. La búsqueda avanzada, en cambio, aportará datos más específicos combinando diferentes criterios de consulta. Por ejemplo, se podrán obtener todas las respuestas que recoge un atlas sobre un campo semántico y un punto de encuesta concreto. Esto supondrá poder estudiar el léxico de una localidad y, a su vez, servirá también para motivar nuevas encuestas que permitan investigar la mortandad y la vitalidad léxicas de un lugar o una región. Este es el objetivo del proyecto [VitaLex](#) para la zona de la Alpujarra (Andalucía); cuyo estudio se ha basado en el análisis de las respuestas léxicas de 10 puntos de encuesta del ALEA en los que se han vuelto a realizar entrevistas con el fin de determinar los cambios que se han producido en cincuenta años. Por todo ello, creemos que las posibilidades de explotación del corpus contribuirán desde diferentes perspectivas al estudio de la variación y las variedades de la península ibérica.

La presentación de los registros se ofrecerá en formato de lista, de forma similar a como se visualizan los datos en algunos corpus textuales ([CORPES XXI](#), [CREA](#), [CORDE](#), [CNDHE](#)) o en otras herramientas geolingüísticas como los [Índices Léxicos do Atlas Lingüístico Galego](#). El objetivo no es, por tanto, ofrecer los datos cartografiados y sustituir el mapa de los atlas, sino recopilarlos para que sean manejables y contrastables de forma más rápida y sencilla, ya que, de acuerdo con Heap (2002: 18):

aunque pueda parecer contradictorio, muchas veces incluso resulta preferible, ya que publicar los datos como listados de formas correspondientes a cada punto de encuesta agiliza su consulta. De esta manera, [...] se facilita la consulta por los

especialistas, que muchas veces se ve hasta dificultada por la necesidad de examinar un elevado número de mapas de gran tamaño para volver a establecer la lista de formas (de unas localidades específicas) que se requieren para un trabajo geolingüístico. Claro está, sin embargo, que para fines de la geografía lingüística es útil a veces poder visualizar la disposición espacial de las formas (o, por lo menos, de símbolos que sintetizan las formas), pero incluso para la eventual cartografía electrónica es preferible partir de una base de datos digitalizada. Tal es el procedimiento seguido en el caso del *Survey of English Dialects*, que fue publicado originalmente como listado de formas [...]. Los elevados costes de la edición de los atlas lingüísticos (impresión especializada en formatos poco usuales) también se pueden evitar utilizando las nuevas tecnologías.

3.2. Aportaciones del corpus al estudio de la lengua: un ejemplo

Desde que se idearon y publicaron los primeros atlas lingüísticos en el siglo XIX, se ha evidenciado que constituyen una fuente de información imprescindible —aunque sea cuestionable su representatividad y la imagen lingüística que proyectan (Caravedo 2005, Álvarez Pérez 2016)—, no solo por los datos que contienen, sino también por la forma en que se representan y organizan (Gilliéron 1918; Jaberg 1959; García Mouton 1990). Joan Coromines (*DECH*), como se ha indicado en numerosas ocasiones (Blecua y Clavería 1999; Prat 2006; García Mouton 2016), señalaba en la introducción de su diccionario etimológico —muy consciente de la importancia de la geolocalización de las palabras al describir la historia de una lengua— que a falta de atlas lingüísticos, había precisado la localización de las palabras que no son de uso general:

era preciso localizar sistemáticamente las palabras, siempre que no sean de uso bien general, y a falta del Atlas Lingüístico (o los Atlas) que esperamos, no he vacilado en dar los informes sueltos que poseo, aun cuando yo mismo los crea sólo parciales, y aun cuando no parezcan útiles para el estudio etimológico; saber que una palabra se emplea en Almería, en Aragón o en la Argentina, p. ej., aunque su área sea en realidad mucho más extensa, nunca podrá desorientarnos (si tenemos conciencia del carácter incompleto de los datos) y no faltarán ocasiones, hoy o en el futuro, en que el dato pueda aclarar un pormenor (*DECH, Introducción*).

El valor que Coromines otorgaba a los atlas para complementar la historia y la etimología del léxico se puede advertir no solo en esta parte de la introducción de la obra, también se evidencia en otros apartados. Por un lado, en la lista de indicaciones bibliográficas aparecen referencias a varios atlas lingüísticos⁸:

- a) G. Bottiglioni, *Atlante Linguistico-Etnografico Italiano della Corsica*, 8 vols., Pisa 1933-9.
- b) *AIS = Sprach- und Sachatlas Italiens und der Südschweiz*, von K. Jaberg und J. Jud; die Mundartenaufnahmen wurden durchgeführt von P. Scheuermeier, G. Rohlf's und M. L. Wagner; 8 vols.; Zofingen, 1928 ss.
- c) *ALC = Atlas Lingüístic de Catalunya*, p. p. A. Griera, vols. 1-4, B. 1923 ss.

⁸ Se indican las informaciones según aparecen en el diccionario.

- d) *ALF* = *Atlas Linguistique de la France*, p p. J. Gilliéron et E. Edmont, P. 1903-10.
e) P. Gardette, *Atlas Linguistique et Ethnographique du Lyonnais*, Lión, 1950 ss.

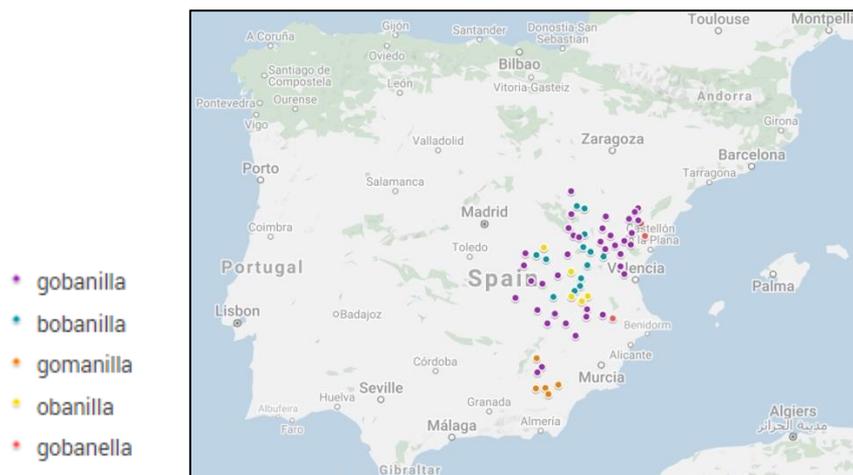
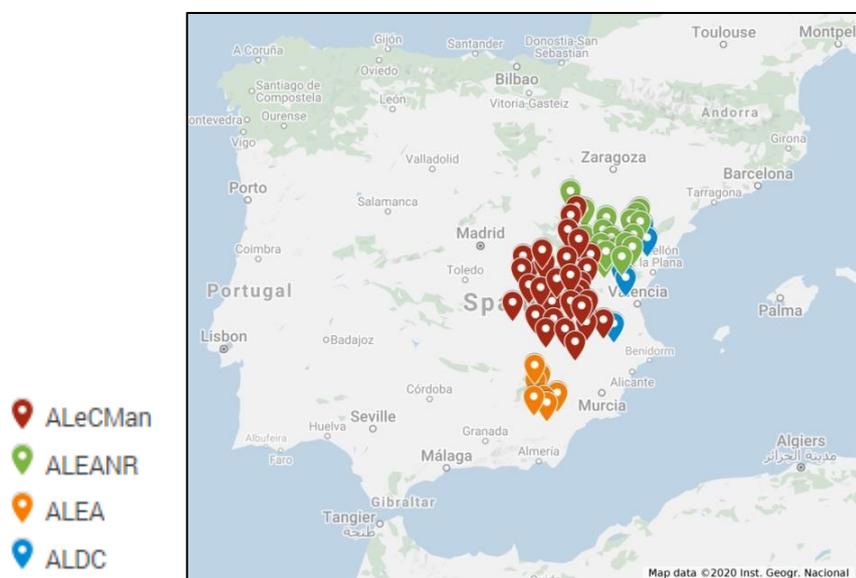
Según la consulta que hemos realizado a partir de la búsqueda abierta de la versión electrónica de la obra —que permite buscar cualquier palabra en cualquiera de los apartados (etimología, desarrollo crítico, documentación, derivados, compuestos, notas)—, no todos estos se citan en el cuerpo de la obra. De los cinco, el último no lo hemos hallado en ninguna de las entradas, aunque sí aparecen otras obras de Gardette mencionadas (p. ej. *Pouèmes dou Païsan*). Por otro lado, gracias al mismo tipo de consulta, se comprueba que para la redacción de la obra se manejaron más atlas de los que se citan en las indicaciones bibliográficas. En concreto, se han hallado menciones a diez atlas lingüísticos de pequeño y gran dominio de diferentes lenguas románicas en cientos de entradas: el *Atlas Lingüístic de Catalunya* de Antoni Griera (s. v. *ballico, bandullo, bausán, birlo, boñiga, cabrio, caldo, chaleco, chivo, eraje, gavilán, grifo, guadaña, limeta, magaña, manga, mendrugo, muñeca, patarata, zueco*); el *Petit Atlas linguistique d'une région des Landes, contribution à la dialectologie gasconne* de Georges Millardet que se cita como *Atlas des Landes* o simplemente por el apellido del autor (s. v. *apañar, gayuba, morueco, púa, sibil vómer, zanca*); el *Atlas Linguistique de la Lozère* de Rudolf Hallig (s. v. *encinta*, nota 1); el *Atlas Linguistico Italiano* de Gino Bottiglioni (s. v. *malecón, morir, serba*); el *Atlas Linguistique de la France*, que se cita como *ALF* (s. v. *alfalfa, andén, antosta, argamasa, ballico, barajar, bausán, gayuba*, etc.); el *Sprach-und Sachatlas Italiens und der Südschweiz*, que se cita como *AIS* (s. v. *andén, atempa, caracol, cogote, cundir, charlar, deslizar, escullirse, grulla, helecho, malecón, manteca, máscara, medio, morir, nata, nuez*, etc.); el *Atlante Linguistico Mediterraneo* que se cita mediante materiales no publicados como *Boll. Atl. Ling. Medit.* o como *Atlante Lingüístico Mediterráneo. Saggio delle Carte* (s. v. *avería, róbalo*); el *Atlante Linguistico della Corsica* de Gino Bottiglioni (s. v. *malecón*); el *Atlas Ling. Romîniei* (s. v. *pellizcar*); y el *Atlas Linguistique des Pyrénées orientales* de Henri Guiter (s. v. *tábano*). Entre estos, el que aparece con más recurrencia en las argumentaciones etimológicas es, con mucha diferencia, el *ALF* (la búsqueda abierta por esta abreviatura arroja 248 ocurrencias, aunque no todas se corresponden con la cita al atlas). Las referencias se emplean tanto para indicar la presencia o la ausencia de un término en los atlas. Las documentaciones geolingüísticas sirvieron no solo para situar en el espacio el uso de una forma sino también para explicar evoluciones fonéticas, documentar relaciones genéticas entre las variedades románicas, justificar el sentido y el uso de una variante, etc.

No es de extrañar que la referencia al *ALF* sea la más frecuente en el *DECH* si se tiene en cuenta que Joan Coromines trabajó con Jules Gilliéron en la década de 1920 mientras preparaba su tesis doctoral sobre vocabulario aranés (Coromines 1931); así lo refleja la correspondencia (fechada en 1926) que mantuvo con sus padres, según Rusinés (1999: 174-175). El número de referencias al *ALF* contrasta, sin embargo, con la ausencia de menciones a los atlas regionales, ya que, cuando se empezó a publicar la segunda edición de la obra, algunos de ellos ya estaban publicados (*ALEA*, *ALEICan*) o estaban en proceso (*ALEANR*). Germà Colón (1994: 600 *apud* Rusinés 1999: 175) critica esta falta de aprovechamiento que de los atlas regionales hizo Coromines en la edición segunda del diccionario etimológico: «El Sr. Corominas no ha abierto un atlas, primero porque al hacer la edición de 1954 apenas los había en el dominio español; y luego,

cuando ya empezó a haberlos, porque los ha desdeñado, y punto. Es ese en su obra un hueco metodológico difícil de colmar».

Las posibilidades de explotación de los mapas geolingüísticos tradicionales no se limitan a permitir complementar una descripción etimológica o la historia de una voz, aunque esto constituya uno de sus principales atributos (Pop 1950; García Mouton 1990; Blecua y Clavería 1999). Desde que se idearon, se han publicado cientos de trabajos de investigación en los que se demuestra de qué modo pueden contribuir estos materiales al análisis lingüístico, cultural y etnográfico tanto desde una perspectiva particular como general. Los trabajos revelan que la importancia de los datos almacenados en ellos puede servir para caracterizar los dialectos (Montero Curiel 2004) y trazar áreas dialectales (Aurrekoetxea 2019); estudiar la lengua oral (García Mouton 1997); investigar la unidad y la diversidad lingüísticas (Julià 2009) y culturales (Viereck 2000); examinar la variación (y su distribución en el espacio y el tiempo) desde el léxico (Veny 2000; Álvarez Pérez 2008; Perea y Colón 2010; Negro 2012), la semántica (Salvador 1965), la fonética (Llorente 1962; Molina 2013), la morfología (Uritani Berrueta 1985; Rosales Solís 2016), la sintaxis (Mouzo y Pérez 2015; Castillo y Octavio de Toledo 2016; Massanell 2017b) o la paremiología (Gargallo 2011); determinar el cambio lingüístico (Molina 2006; Herrgen 2010; Del Barrio 2018); estudiar la formación histórica de la variedad estándar (Fernández-Ordóñez 2011); llevar a cabo estudios etnográficos de las comunidades lingüísticas (Munévar y Bernal 2018); analizar el contacto de lenguas (Alvar 1976; Martines 2012); establecer los procedimientos de creación léxica más frecuentes en la lengua popular (Fuster 1996); proporcionar información sobre la motivación lingüística (Alinei 2002); examinar la productividad de los procesos de formación de palabras (García González 2018); mejorar y completar los datos lexicográficos (Fernández Morell 2015); y, entre otros, contribuir al estudio de las relaciones genéticas entre variedades lingüísticas (Tagliavini 1949).

El uso de *CORPAT* puede servir para ampliar esta lista de líneas de investigación, para arrojar luz sobre asuntos que no han sido tenidos en cuenta en la dialectología tradicional (p. ej. el estudio de las respuestas múltiples o polimorfismo, Aurrekoetxea 2019: 35) y para permitir rastrear parte de la extensión geográfica de las voces que no documentan ni los corpus ni los diccionarios. En anteriores trabajos se mencionó, como ejemplo de este último asunto, el caso de *gobanilla* ‘muñeca de la mano’ (y sus variantes *gobanella* o *govanella*, *gomanilla*, *bobanilla*, *obanilla*) que aparece en cuatro atlas de la península ibérica (mapa 98 del *ALDC* para el catalán y los mapas 1267 del *ALEA*, *982 del *ALEANR* y 334 del *ALeCMan*) con más de sesenta ocurrencias (figuras 13 y 14), la mayoría de las cuales se recogen en el *ALEANR* (17) y el *ALeCMan* (35).

Figura 14. Ocurrencias de *gobanilla* y sus variantes en los atlas regionalesFigura 15. Mapa de las ocurrencias de *gobanilla* y sus variantes distribuidas por atlas lingüístico

Los atlas muestran claramente una localización oriental para esta forma de designar la muñeca de la mano, pero ninguna de las variantes recogidas en el mapa de la figura 14 aparece documentada en los corpus y repertorios lexicográficos generales del español (*CORDE*, *CNDHE*, *CREA*, *DLE*, *NTLLE*), ni tampoco en el *Corpus Informatitzat del Català Antic* (*CICA*). Esta ausencia sorprende si se tiene en cuenta que no es extraño poder rastrear dialectalismos orientales en el *CORDE* (Martines 2012). Existen, no obstante, algunos testimonios en estudios dialectales (Chacón 1981; Llorente 1985; Gargallo 1987: 295; *DCVB*), en diccionarios etimológicos (*DECH*, *DECat*) y en el *Corpus Informatitzat del Valencià* (*CIVAL*), que coinciden con los atlas al situar las formas siempre en la zona oriental de la Península (Granada, Albacete, Cuenca, Murcia, Alicante, Valencia, Teruel y Castellón). Por ello, no parece desacertado, gracias al dibujo que permiten trazar los atlas de los que disponemos, suponer que se trata de una voz que se desplaza de norte a sur debido a los movimientos de población que se producen en la Edad Media. Su uso medieval está documentado en el diccionario catalán-latín *Thesaurus puerilis* de Onofre Pou (1575) —donde el gerundense afinado

en valencia «copiaba frases enteras y palabras que oía a sus estudiantes» (Casanova 2011: 221) entre las que se pueden encontrar bastantes aragonesismos y castellanismos—. Además, según las nueve ocurrencias del *Corpus Informatizat del Valencià (CIVAL)*, la forma está viva todavía en la zona septentrional del valenciano.

Todos estos datos, junto con la ausencia de las formas en los documentos y corpus mencionados, podrían indicar que se trata de un aragonesismo, lo que se justificaría por la elevada presencia de pobladores de posible ascendencia aragonesa desde los primeros siglos de la conquista hasta la fundación del reino de Valencia (Martines 2012: 127), e iría en la línea de algunas investigaciones en las que se señala la influencia del aragonés en las variedades del sur de Cataluña y en el valenciano (Veny 1991; Casanovas 2011). Más difícil parece determinar la etimología de estas variantes a partir de las diferentes propuestas de Joan Coromines que, por un lado, en el diccionario etimológico del español, vincula *gobanella* con el mozárabe *qobtel*, *-tal*, ‘medida desde el codo a la muñeca’ (DCECH s. v. *muñeca*); y, por el otro, en el diccionario etimológico catalán se refiere a la forma *MUNNONELLA, que sería derivado diminutivo del prerromano *MUNNONE, posible origen la voz *munyó* ‘muñón’ (DECat, s. v. *govanella*) sin descartar un posible origen mozárabe distinto al que propuso en el diccionario castellano: «L’alternativa més fundada a aquesta idea em semblaria admetre una evolució mossàrab de GUBERNACULA (pl. gubernaculum) ‘govern’, ‘governalls’» (DECat, s. v. *govanella*).

4. CONCLUSIÓN

Hasta el momento, los datos de los atlas regionales se han aprovechado parcialmente gracias a que algunos investigadores interesados en un tema lingüístico concreto (de carácter léxico, fonético, morfológico o sintáctico) se han acercado a ellos, han vaciado los mapas y los han estudiado en contraste con otras fuentes de información. Por ello, la bibliografía en la que se exponen los resultados de estas indagaciones es el único medio del que disponemos actualmente para determinar cuál es el provecho de estos materiales. La publicación de *CORPAT* facilitará el acceso a los datos y aumentará las posibilidades de explotarlos en conjunto desde múltiples perspectivas para extraerles el máximo rendimiento. Esto, en definitiva, se corresponde con lo que se pretendió al inicio de la automatización de los atlas regionales del español:

Los Atlas Lingüísticos habrán dejado de ser ese monstruoso fantasma que abrumba por su magnitud. Ahora ya se podrán usar como diccionario normal, como repertorio de sufijos, desinencias, etc., como conjunto para estudiar campos semánticos, como elementos fundamentales de estadística lingüística y otros etcéteras no poco extensos. Unase [sic.] esto a las virtudes clásicas de la geografía lingüística (distribución espacial ordenada, visión simultánea de los fenómenos, ordenación geográfica, transcripción rigurosa, recogida *in situ*, etc.) (Alvar y Nuño 1981: 371).

En esencia, atesorar la información de todos los atlas regionales en bases de datos supondrá no solo que se puedan usar los materiales que guardan los mapas, sino que también servirá para preservarlos y evitar que se pierdan. Paradójicamente, algunos de los objetivos de la geografía lingüística del siglo XXI coinciden en parte con

una de las principales razones por la que se crearon los atlas en el siglo XIX (Coseriu 1977: 118): evitar que los datos que se copiaron puedan perderse y ponerlos a disposición de la sociedad y de la comunidad académica para que puedan ser estudiados (Heap 2002; Pato 2009 en ALPR; Sousa 2017).

REFERENCIAS BIBLIOGRÁFICAS

- ADIM = García Mouton, Pilar e Isabel Molina Martos (2015): *Atlas Dialectal de Madrid*. Madrid: CSIC. <http://adim.cchs.csic.es/es> [Consulta: 15/03/2020].
- AIS = Jaberg, Karl y Jakob Jud (1928-1940): *Sprach-und Sachatlas Italiens und der Südschweiz*. Zofingen: Gedruckt mit Unterstützung der Gesellschaft für Wissenschaftliche Forschung an der Universität Zurich und privater Freunde des Werkes von der Verlagsanstalt Ringier & Co., 8 vols.
- ALCYL = Alvar, Manuel (1999): *Atlas Lingüístico de Castilla y León*, Valladolid: Junta de Castilla y León / Consejería de Educación y Cultura, 3 vols.
- ALDC = Veny, Joan y Lúdia Pons (2001-2018): *Atles lingüístic del domini català*. Barcelona: Institut d'Estudis Catalans, 9 vols. <http://aldc.espais.iec.cat/mapes/> [Consulta: 15/03/2020].
- ALEA = Alvar, Manuel (1961-1973): *Atlas lingüístico y etnográfico de Andalucía*. Granada: Universidad de Granada/CSIC, 6 vols. [Con la colaboración de Antonio Llorente y Gregorio Salvador].
- ALEANR = Alvar, Manuel, Antonio Llorente y Tomás Buesa (1979-1983): *Atlas lingüístico y etnográfico de Aragón, Navarra y Rioja*: Madrid: La Muralla, 12 vols.
- ALEC = Instituto Caro y Cuervo (2018): *Atlas Lingüístico-Etnográfico de Colombia*. <http://alec.caroycuervo.gov.co> [Consulta: 15/03/2020].
- ALECan = Alvar, Manuel (1995): *Atlas lingüístico y etnográfico de Cantabria*. Madrid: Arco/Libros, 2 vols.
- ALECh = Wagner, Claudio (2014-2015): *Atlas Lingüístico y etnográfico de Chile*. <http://www.atlaslinguistico.cl/> [Consulta: 29/10/2020].
- ALECMAN = García Mouton, Pilar y Francisco Moreno Fernández (2003): *Atlas lingüístico y etnográfico de Castilla-La Mancha*. Alcalá de Henares: Universidad de Alcalá de Henares <https://www.linguas.net/alecman/> [Consulta: 15/03/2020].
- ALEICan = Alvar, Manuel (1975-1978): *Atlas lingüístico y etnográfico de las Islas Canarias*. Madrid: La Muralla, 3 vols.
- ALF = Gilliéron, Jules y Edmond Edmont (1902-1910): *Atlas Linguistique de la France*. Paris: Honoré Champion, 12 vols.
- ALPI = García Mouton, Pilar (coord.), Inés Fernández-Ordóñez, David Heap, María Pilar Perea, João Saramago y Xulio Sousa (2016) ALPI-CSIC, edición digital de Navarro Tomás, Tomás (dir.): *Atlas Lingüístico de la Península Ibérica*. Madrid: CSIC. <http://www.alpi.csic.es/> [Consulta: 15/03/2020].
- ALPR = Navarro Tomás, Tomás (1948): «Atlas Lingüístico de Puerto Rico», en *El español en Puerto Rico: Contribución a la geografía lingüística hispanoamericana*, Puerto Rico: Río Piedras. <http://www.alpr.info/> [Consulta: 29/10/2020]
- ALINEI, Mario (2002): «Il ruolo della motivazione nel lessico», en Rosario Álvarez Blanco, Francisco Dubert García y Xulio Sousa Fernández (eds.): *Dialectología e Léxico*. Santiago de Compostela: Consello da Cultura Galega/Instituto da Lingua Galega, pp. 15-28.
- ALVAR, Manuel (1976): «Un problema de lenguas en contacto: la frontera catalano-aragonesa», *Archivo de filología aragonesa*, 18-19, pp. 23-38.
- ALVAR, Manuel (1982): «Atlas lingüísticos y diccionarios», *Lingüística Española Actual*, IV/2, pp. 253-323.

- ALVAR, Manuel (1991): «El Atlas Lingüístico y Etnográfico de la provincia de Santander», en Manuel Alvar (dir.): *Estudios de geografía lingüística*, Madrid: Paraninfo, pp. 349-378.
- ALVAR, Manuel y M.ª Pilar NUÑO (1981): «Un ejemplo de atlas lingüístico automatizado: el ALES», *Lingüística Española Actual*, III, pp. 359-370.
- ALVAR, Manuel y Manuel VERDEJO (1978): «Automatización de atlas lingüísticos», *Revista de Dialectología y Tradiciones Populares*, XXXIV, pp. 23-48.
- ÁLVAREZ PÉREZ, Xosé Afonso (2008): *O léxico da vaca. Nomes basados no físico*. Santiago de Compostela: Universidad de Santiago de Compostela. Tesis doctoral dirigida por Manuel González González. <http://hdl.handle.net/10347/2438> [Consulta: 15/03/2020].
- ÁLVAREZ PÉREZ, Xosé Afonso (2016): «Consideraciones sobre la imagen lingüística transmitida por los atlas tradicionales: el caso del *Atlas Lingüístico de la Península Ibérica* en Portugal», *Géolinguistique*, 16, pp. 131-160. <http://journals.openedition.org/geolinguistique/528> [Consulta: 15/06/2020].
- ARIAS, Alberto Manuel, Mercedes DE LA TORRE y María Isabel FIJO (2013): *ICTIOTERM. Base de datos terminológica y de identificación de especies pesqueras de las costas de Andalucía*. CSIC. <http://www.ictioterm.es/> [Consulta: 15/03/2020].
- ARIAS, Laura y Ángel J. GALLEGRO BARTOLOMÉ (2019): «El Atlas Dialectal de Barcelona Una herramienta para explorar la variación lingüística», *Congreso ALFALito 2019 (Queens College, Nueva York)*. http://adiba.cat/wp-content/uploads/2019/09/ADIBA_Alfalito-2_compressed.pdf [Consulta: 15/06/2020].
- ASinEs = *Atlas Sintáctico del Español*. <http://www.asines.org> [Consulta: 15/03/2020].
- AURREKOETXEA, Gotzon (2006 [2002]): «El Atlas Lingüístico Vasco (EHHA)», M. Aurnague, M. Roché (dir.): *Homenage à Jacques Allières-1. Domaines Basque et Pyrénéen*. Biarritz: Editions Atlantica, pp. 63-71. https://artxiker.ccsd.cnrs.fr/file/index/docid/71353/filename/El_Atlas_Ling_Vasco.pdf [Consulta: 15/06/2020].
- AURREKOETXEA, Gotzon (2019): «Sobre el valor de la dialectometría en la delimitación de las distancias lingüísticas», *GLOSEMA. Revista Asturiana de Lingüística*, I, pp. 19-39.
- BLECUA, José Manuel y Gloria CLAVERÍA (1999): «La lexicografía castellana, antes y después de Joan Coromines», en Joan Solà (ed.): *L'obra de Joan Coromines. Cicle d'estudi i homenatge*, Sabadell: Fundació Caixa de Sabadell-Aula de Ciència i Cultura, pp. 29-43.
- BONILLA, Jonathan E. y Julio A. BERNAL CHÁVEZ (2020): «Modelamiento de una base de datos espacial para el *Atlas Lingüístico-Etnográfico de Colombia*», *Revista Signos. Estudios de Lingüística*, 53, pp. 346-368. <http://www.revistasignos.cl/index.php/signos/article/view/106/202> [Consulta: 15/06/2020].
- BOSQUE, Isabel del, Carlos FERNÁNDEZ, Lourdes MARTÍN-FORERO y Esther PÉREZ (2012): *Los sistemas de información geográfica y la investigación en ciencias humanas y sociales*. Madrid: Confederación Española de Centros de Estudios Locales. <http://digital.csic.es/bitstream/10261/64940/1/Los%20SIG%20y%20la%20Investigacion%20en%20Ciencias%20Humanas%20y%20Sociales.pdf> [Consulta: 15/06/2020].
- BOUZOUITA, Miriam, Mónica CASTILLO y Enrique PATO (2018): «*Dialectos del español*. Una nueva aplicación para conocer la variación actual y el cambio en las variedades del español», *Dialectología: revista electrónica*, 20, pp. 61-83. <https://www.raco.cat/index.php/Dialectologia/article/view/334686> [Consulta: 15/06/2020].
- BROWN, Cecil H. (2013): «Hand and Arm», en Dryer, Matthew S. y Haspelmath, Martin (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/129> [Consulta: 15/08/2020].

- BUZAI, Gustavo D. (2004): «Geografía y tecnologías digitales del siglo XXI: una aproximación a las nuevas visiones del mundo y sus impactos científico-tecnológicos», *Scripta Nova. Revista electrónica de geografía y ciencias sociales*, VIII/170 (58), <http://www.ub.edu/geocrit/sn/sn-170-58.htm> [Consulta: 15/03/2020].
- CARAVEDO, Rocío (2005): «Sobre el corpus en la dialectología actual», en Mario Lozano Ramírez (ed.): *Homenaje a José Joaquín Montes Giraldo*. Bogotá: Instituto Caro y Cuervo, pp. 94-107.
- CASANOVA, Emili (2011): «Influencia histórica del aragonés sobre el valenciano», *Archivo de Filología Aragonesa (AFA)*, 67, pp. 201-235.
- CASTAÑER MARTÍN, Rosa M.^a (2009): «El atlas lingüístico como fuente lexicográfica. Del ALEANR al DDEAR», *Archivo de filología aragonesa (AFA)*, 65, pp. 139-164. <https://ifc.dpz.es/recursos/publicaciones/29/56/06castaner.pdf> [Consulta: 15/06/2020].
- CASTILLO, Mónica y Álvaro S. OCTAVIO DE TOLEDO Y HUERTA (2016): «*Habemos muchos que hablamos español*: distribución e historia de la concordancia existencial en primera persona de plural», en Carlota de Benito y Álvaro S. Octavio de Toledo y Huerta (eds.): *En torno a 'haber'. Construcciones, usos y variación desde el latín hasta la actualidad*. Bern: Peter Lang, pp. 111-168.
- CNDHE = Instituto de Investigación Rafael Lapesa de la Real Academia Española (2013): *Corpus del Nuevo Diccionario Histórico (CDH)*. <http://web.frl.es/CNDHE>. [Consulta: 15/06/2020].
- CERRUDO, Alba, Ángel J. GALLEGO, Anna PINEDA y Francesc ROCA (2015): «ASinEs: Prolegómenos de un atlas de la variación sintáctica del español», *Linguamática*, 7/2, pp. 59-69. <https://linguamatica.com/index.php/linguamatica/article/view/V7N2.5/342> [Consulta: 15/06/2020].
- CHACÓN, Carmen (2016): *Análisis demolingüístico del léxico hispánico: Estudio aplicado a las "naciones específicas" del Plan Curricular del Instituto Cervantes*. Madrid: Universidad Nacional de Educación a Distancia. Tesis doctoral dirigida por M.^a Antonieta Andión Herrero. http://e-spacio.uned.es/fez/eserv/tesisuned:ED-Pg-Filologia-Chacon/CHACON_GARCIA_Carmen_Tesis.pdf [Consulta: 15/06/2020].
- CICA = Torruella Casañas, Joan, Manel Pérez Saldanya y Josep Martines Peres (coords.): *Corpus del català antic*. <http://cica.cat/index.php> [Consulta: 15/06/2020].
- CIVAL = Acadèmia Valenciana de la Llengua. *Corpus Informatitzat del Valencià*. <http://cival.avl.gva.es/cival/buscador.jsp> [Consulta: 15/06/2020].
- COLON, Germà (1994): «Sobre los estudios de etimología española», en *Actas del congreso de la lengua española*, Madrid: Instituto Cervantes, pp. 597-610.
- CORDE = Real Academia Española: Banco de datos (CORDE) [en línea], *Corpus diacrónico del español*. <http://www.rae.es>. [Consulta: 15/06/2020].
- COROMINES, Joan (1931): *Vocabulario aranés. Tesis doctoral*. Barcelona: Imprenta de la Casa de la Caridad.
- CORPES XXI = Real Academia Española: Banco de datos (CORPES XXI) [en línea], *Corpus del Español del Siglo XXI (CORPES)*. <http://www.rae.es> [Consulta: 29/10/2020].
- COSER = Fernández Ordóñez, Inés (dir.) (2005-): *Corpus Oral y Sonoro del Español Rural*. www.corpusrural.es [Consulta: 15/06/2020].
- COSERIU, Eugenio (1977): «La geografía lingüística», en *El hombre y su lenguaje. Estudios de teoría y metodología lingüística*. Madrid: Gredos, pp. 103-158.
- CREA = Real Academia Española: Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual*. <http://www.rae.es> [Consulta: 15/06/2020].
- DAUZAT, Albert (1944 [1922]): *La géographie linguistique*. Paris: Falammariion.
- DCVB = Alcover, Antoni M.^a y Francesc de Borja Moll (2002 [1962]): *Diccionari català-valencià-balear*. Barcelona: Institut d'Estudis Catalans/Editorial Moll. <http://dcvb.iecat.net/>. [Consulta: 15/06/2020].

- DECat = Joan Coromines (1980-1991): *Diccionari Etimològic i Complementari de la Llengua Catalana*. Barcelona: Curial. [Con la colaboración de Joseph Gulsoy y Max Cahner].
- DECH = Joan Coromines y José Antonio Pascual (1980-1991): *Diccionario Crítico Etimológico Castellano e Hispánico*. Madrid: Gredos. Edición digital en CD-ROM (2012).
- DEL BARRIO DE LA ROSA, Florencio (2018): *Espacio variacional y cambio lingüístico en español*, Madrid: Visor.
- DLE = Real Academia Española (2014): *Diccionario de la lengua española*, Madrid: Espasa Calpe. <https://dle.rae.es/> [Consulta: 15/03/2020].
- ESTRADA ARRÁEZ, Ana y Carlota DE BENITO MORENO (2016a): «Variación en las redes sociales: datos twilectales», *Revista Internacional de Lingüística Iberoamericana*, 28/2, pp. 77-11.
- ESTRADA ARRÁEZ, Ana y Carlota DE BENITO MORENO (2016b): «Introducción. Si Alvar levantara la cabeza: cómo tratar con informantes virtuales», *Revista Internacional de Lingüística Iberoamericana*, 28, pp. 7-17.
- ESTRADA ARRÁEZ, Ana y Carlota DE BENITO MORENO (2018): «Aproximación metodológica al estudio de la variación lingüística en las interacciones digitales», *REDD: Revista de estudios del discurso digital*, 1, pp. 74-122.
- FERNÁNDEZ CAMPOS, Javier Orlando (2017): *Diseño e implementación de un sistema de información geográfica en línea para el Atlas lingüístico-etnográfico de Colombia, su manual y suplemento*. Bogotá: Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería. Trabajo de Grado Modalidad monografía para optar por el título de ingeniero catastral y geodesta. <http://repository.udistrital.edu.co/handle/11349/7730> [Consulta: 15/06/2020].
- FERNÁNDEZ MORELL, M.^a Lourdes (2015): «Las marcas diatópicas del DRAE y los atlas lingüísticos españoles. Correspondencia de áreas y problemas derivados de la marcación diatópica», *Tonos digital. Revista de Estudios filológicos*, 29, pp. 1-36.
- FERNÁNDEZ MORELL, M.^a Lourdes (en prensa): «Los nombres de animales y vegetales como patrimonio lingüístico alpujarreño a partir de los datos del proyecto VitaLex» <http://www.proyectovitalex.es/pdf/articulos/2019-05-publicaciones.pdf> [Consulta: 15/06/2020].
- FERNÁNDEZ-ORDÓÑEZ, Inés (2011): *La lengua de Castilla y la formación del español*. Discurso leído el día 13 de febrero de 2011 en su recepción pública. Madrid: Real Academia.
- FERNÁNDEZ SEVILLA, Julio (1974): «Lexicografía y geografía lingüística», *Problemas de la lexicografía actual*. Bogotá: Instituto Caro y Cuervo, pp. 81-113.
- FUSTER, M.^a Teresa (1996): «Voces de creación metafórica sobre el maíz y el trigo en el Atlas Lingüístico y Etnográfico de Aragón, Navarra y Rioja», *Estudios de Lingüística de la Universidad de Alicante (ELUA)*, 1, pp. 139-147.
- GARCÍA GONZÁLEZ, Elisa (2018): «El Atlas Lingüístico de Cuba: un proyecto finalmente concluido», *Universidad de La Habana*, 286, pp. 110-120. <http://www.revuh.uh.cu/index.php/UH/article/view/172/80> [Consulta: 15/06/2020].
- GARCÍA MOUTON, Pilar (1990): «El estudio léxico en los mapas lingüísticos», en Francisco Moreno Fernández (recop.), *Estudios sobre variación lingüística*. Salamanca: Universidad de Alcalá de Henares, pp. 27-75.
- GARCÍA MOUTON, Pilar (1997): «El hablante en los textos orales», en Luis Díaz G. Viana y Matilde Fernández Montes (coords.), *Entre la palabra y el texto. Problemas en la interpretación de fuentes orales y escritas. Curso de Etnología Española "Julio Caro Baroja" XVI Edición*. Madrid: CSIC, pp. 233-244.
- GARCÍA MOUTON, Pilar (2009): «La investigación geolingüística española en la actualidad», en Dolores Corbella y Josefa Dorta (eds.), *La investigación dialectológica en la actualidad*, Santa Cruz de Tenerife: Agencia Canaria de Investigación, Innovación y Sociedad de la Información. Gobierno de Canarias, pp. 333-346.

- GARCÍA MOUTON, Pilar (2010a): «Si el *DRAE* utilizase los atlas lingüísticos...: un ejemplo aragonés», en Rosa M.^a Castañer Martín y Vicente Lagüéns Gracia (eds.), *De moneda nunca usada. Estudios filológicos dedicados a José M.^a Enguita Utrilla*. Zaragoza: Inst. «Fernando el Católico», pp. 271-282.
- GARCÍA MOUTON, Pilar (2010b): «El procesamiento informático de los materiales del *Atlas de la Península Ibérica* de Tomás Navarro Tomás», en Gotzon Aurrekoetxea y José Luis Ormaetxea (eds.), *Tools for linguistic variation*. Bilbao/Bilbo: Universidad del País Vasco/Euskal Herriko Unibertsitatea, pp. 167-174.
- GARCÍA MOUTON, Pilar (2016a): «Claudio Wagner, *El Atlas Lingüístico y Etnográfico de Chile (ALECh) por regiones*», *Géolinguistique*, 16, pp. 255-260. <https://journals.openedition.org/geolinguistique/pdf/489> [Consulta: 29/10/2020].
- GARCÍA MOUTON, Pilar (2016b): «Dialectología y geografía lingüística», en Javier Gutiérrez Rexach (coord.), *Enciclopedia Lingüística Hispánica*. New York: Rotudledge, vol. I, pp. 30-40.
- GARCÍA MOUTON, Pilar (2016c): «Corominas tenía razón: *jamila* no *jámila*» en Mariano Quirós (ed.), *Etimología e historia en el léxico del español. Estudios ofrecidos a José Antonio Pascual (Magister bonus et sapiens)*. Madrid/Frankfurt: Iberoamericana/Vervuert, pp. 293-302.
- GARCÍA MOUTON, Pilar (2017): «El *Atlas Lingüístico de la Península Ibérica (ALPI)* en línea. Geolingüística a la carta», *Estudis romànics*, 39, 335-343.
- GARGALLO, José Enrique (1987): *Una encrucijada lingüística entre Aragón, Valencia y Castilla: el rincón de Ademuz*. Barcelona: Universitat de Barcelona. Tesis doctoral.
- GARGALLO, José Enrique (2011): «Refráns meteorológicos e atlas románicos, paremiología e territorio», *Estudos de Lingüística Galega*, 3, pp. 31-50.
- GDLE = Ignacio Bosque y Violeta Demonte (1999): *Gramática Descriptiva de la Lengua Española*, Madrid: Espasa Calpe.
- GILLIÉRON, Jules (1918): *Généalogie des mots qui désignent l'abeille d'après l'Atlas Linguistique de la France*. Paris: Librairie Ancienne Honoré Champion/Éditeur Édouard Champion.
- GIRNTH, Heiko (2010): «Mapping language data», en Alfred Lameli, Roland Kehrein y Stefan Rabanus (eds.), *Language and space. An international handbook of linguistic variation. Language Mapping*. Berlin/New York: Mouton De Gruyter, vol. II, pp. 98-121.
- GOEBL, Hans (2010): «Introducción a los problemas y métodos según los principios de la escuela dialectométrica de Salzburgo (con ejemplos sacados del “Atlante Italo-svizzero”, AIS)», Gotzon Aurrekoetxea Olabarri y José Luis Ormaetxea Lasaga (coords.), *Tools for linguistic variation*, Bilbao/Bilbo: Universidad del País Vasco/Euskal Herriko Unibertsitatea, pp. 3-40.
- HEAP, David (2002): «Segunda noticia histórica del *ALPI*», *Revista de Filología Española*, LXXXII, pp. 5-19.
- HOCH, Shawn C. y HAYES, James J. (2010): «Geolinguistics: The Incorporation of Geographic Information Systems and Science», *The Geographical Bulletin*, 51, pp. 23-36.
- HERRGEN, Joachim (2010): «The digital wenker atlas (www.diwa.info): An online research tool for modern dialectology», *Dialectologia*, I (Special Issue), pp. 89-95.
- JABERG, Karl 1959 ([1908]): *Geografía lingüística. Ensayo de interpretación del “Atlas Lingüístico de Francia”*. Granada: Universidad de Granada. [Traducción española por Antonio Llorente y Manuel Alvar de *Sprachgeographie. Beitrag zum Verständnis des “Atlas Linguistique de la France”*. Aarau: Druck und Verlag von H. R. Sauerländer & Co.]
- JULIÀ LUNA, Carolina (2009): «Los nombres de la pupila en los atlas regionales de la península ibérica», *Lingüística Española Actual*, 31/1, pp. 89-131.
- JULIÀ LUNA, Carolina (2018): «Corrientes teóricas y metodológicas en el estudio del léxico dialectal. Una perspectiva histórica», en María Álvarez de la Granja y Ernesto Xosé González Seoane (eds.), *Léxico dialectal y lexicografía en la Iberorromania*, Madrid/Frankfurt am Main: Iberoamericana/Vervuert, pp. 25-51.

- JULIÀ LUNA, Carolina (2020): «Reseña a Alberto Manuel Arias García y Mercedes de la Torre García (2019): *Ictionimia andaluza. Nombres vernáculos de especies pesqueras del «Mar de Andalucía»*. Madrid: CSIC», *Dialectología et Geolinguística*, 28, pp. 163-172.
- JULIÀ LUNA, Carolina (en prensa): «Los atlas como corpus lingüísticos: más allá de la variación geolectal», *Diversidad: vocablos y voces de las variantes lingüísticas. II Congreso Internacional de Dialectología de Euskaltzaindia* (13-15 de marzo de 2019, Euskaltzaindia, Real Academia de la Lengua Vasca, Bilbao). Frankfurt am Main/Bilbao: Iberoamericana Vervuert/Euskaltzaindia.
- KLEMOLA, Juhani y Mark JONES (1999): «The Leeds Corpus of English Dialects - Project», *Leeds Studies in English*, 30, pp. 17-30. https://ludos.leeds.ac.uk/R/-?func=dbin-jump-full&object_id=124877&silos_library=GEN01 [Consulta: 15/06/2020].
- KOVAL, Santiago (2017): «El signo según Charles Sanders Peirce», *Kubernética. Revista*, <https://www.santiagokoval.com/2017/01/05/el-signo-segun-charles-sanders-peirce/>
- LAMELI, Alfred (2010): «Linguistic atlases – traditional and modern», Peter Auer y Jürgen Schmidt (eds.), *Language and Space. Theories and methods*, vol. 1. Berlin: Mouton De Gruyter, pp. 567-592.
- LIMPER, Juliane, Jeffrey PHEIFF y Anneli WILLIAMS (2020): «REDE SprachGIS: A Geographic Information System for Linguists», en S. Brunn y R. Kehrein (eds.), *Handbook of the Changing World Language Map*. Cham: Springer. https://doi.org/10.1007/978-3-319-73400-2_145-1
- LORENTE, Antonio (1962): «Fonética y fonología andaluzas», *Revista de Filología Española*, XLV, pp. 227-240.
- LORENTE, Antonio (1985): «Coincidencias léxicas entre Andalucía y el Valle del Ebro», *Archivo de Filología Aragonesa*, XXXVI-XXXVII, pp. 347-375.
- MALMBERG, Bertil (1983): *Analyse du langage au xxe siècle. Théories et méthodes*, Paris, Presses Universitaires de France.
- MARTINES, Josep (2012): «Història del lèxic y contacte de llengües. El català i l'aragonès al país valencià a l'edat mitjana: un tast lèxic», en Gloria Clavería, Margarita Freixas, Marta Prat y Joan Torruella (eds.), *Historia del léxico: perspectivas de investigación*. Madrid: Iberoamericana, pp. 127-166.
- MASSANELL, Mar (2017a): «Resseña al Petit Atlas Lingüístic del Domini Català (4 i 5)», *Llengua & Literatura*, 27, pp. 153-157.
- MASSANELL, Mar (2017b): «Los atlas lingüísticos como fuente para estudios de sintaxis dialectal: el caso del Atlas Lingüístic del Domini Català (ALDC)», *Verba. Anuario galego de filoloxía*, 44, pp. 71-98.
- MOLINA, Isabel (2006): «Innovación y difusión del cambio lingüístico en Madrid», *Revista de Filología Española*, LXXXVI/1 (2006) 127-149.
- MOLINA, Isabel (2013): «Yeísmo madrileño y convergencia dialectal campo / ciudad», en Rosario Gómez e Isabel Molina (eds.), *Variación yeísta en el mundo hispánico*. Madrid /Frankfurt am Main, Iberoamericana /Vervuert, pp. 93-110.
- MONTERO BERNAL, Lourdes E. et al. (2018): «Atlas Lingüístico de Cuba (ALCu)», *Revista Anales de la Academia de Ciencias de Cuba*, 8/1, pp. 1-14. <http://www.revistaccuba.sld.cu/index.php/revacc/article/view/413/406> [Consulta: 15/06/2020].
- MONTERO CURIEL, Pilar (2004): «El dialecto leonés y el Atlas Lingüístico de Castilla y León», *Anuario de Estudios Filológicos*, XXVII, pp. 191-205.
- MOSELEY, Christopher (ed.) (2010): *Atlas of the World's Languages in Danger*. Paris: UNESCO Publishing, 3.ª edición. <http://www.unesco.org/culture/en/endangeredlanguages/atlas> [Consulta: 15/06/2020].
- MOUZO, Rosa y Carolina PÉREZ (2015): «Conditional sentences in the ALPI. A Contrastive analyse of the peninsular northwest», *Dialectología*, 5 (Special Issue), pp. 109-140.

- MUNÉVAR, Alejandro y Julio Alexander BERNAL (2018): «Entre unas palabras y unas cosas: la dimensión etnográfica del *Atlas Lingüístico-Etnográfico de Colombia (ALEC)*», *Cuadernos de Lingüística Hispánica*, 31, pp. 19-35.
- NAVARRO CARRASCO, Ana Isabel (1990): «Voces del ALEA no recogidas en el DRAE (1984, 20ª ed.)», *Español actual: Revista de español vivo*, 54, pp. 41-90.
- NAVARRO CARRASCO, Ana Isabel (1992): «Términos del tomo I del ALEICan que no figuran en el DRAE (1984, 20ª ed.)», en Manuel Ariza et al. (coord.), *Actas del II Congreso Internacional de Historia de la Lengua Española*. Madrid: Pabellón de España, vol. 1, pp. 1251-1266.
- NAVARRO CARRASCO, Ana Isabel (1993): «Breve panorama de la dialectología», *Anuario de estudios filológicos*, 16, pp. 309-330.
- NAVARRO CARRASCO, Ana Isabel (1996): «Lexicografía y geografía lingüística. Consideraciones generales», en *El Atlas de Canarias y el Diccionario Académico*. Alicante: Universidad de Alicante, pp. 14-26.
- NAVARRO CARRASCO, Ana Isabel (1998): *Comentario de mapas lingüísticos españoles*. Alicante: Universidad de Alicante.
- NAVARRO CARRASCO, Ana Isabel (2001): «'Yayo' en los diccionarios y en los atlas lingüísticos», *Revista de Filología Española*, 81/1-2, pp. 185-193.
- NEGRO, Marta (2012): «Names of the uvula in the west of the Iberian peninsula: a geolinguistic approach», *Dialectologia*, 3 (Special issue), pp. 111-128.
- NGLE = Real Academia Española (2009): *Nueva Gramática de la Lengua Española*, Madrid: Espasa Calpe.
- NTLLE = Real Academia Española (2001): *Nuevo Tesoro Lexicográfico de la Lengua Española*. Madrid: Espasa. <http://buscon.rae.es/ntlle/SrvltGUILoginNtile> [Consulta: 15/06/2020].
- OROZCO, Manuel (1864): *Geografía de las lenguas y carta etnográfica de México: precedidas de un ensayo de clasificación de las mismas lenguas y de apuntes para las inmigraciones de las tribus*, México: J. M. Andrade y F. Escalante.
- PALDC = Joan Veny y Lídia Pons (2007-2019): *Petit Atlas Lingüístic del Domini Català*. 7 vols. Barcelona: Institut d'Estudis Catalans. <https://aldc.espais.iec.cat/mapes/volum-i/petit-atles-linguistic-del-domini-catala/> [Consulta: 15/06/2020].
- PEREA, M.ª Pilar y Germà COLÓN (2010): «Cronoestratigrafía dialectal», en Maria Iliescu, Heidi Siller-Runggaldier, Paul Danler (eds.), *Actes du XXV^e Congrès International de Linguistique et de Philologie Romanes*, IV. Berlin/New York, De Gruyter, pp. 199-211.
- POP, Sever (1950): *La dialectologie. Aperçu historique et méthodes d'enquêtes linguistiques*, 2 vols. Louvain: Université Catholique de Louvain.
- PRAT Sabater, Marta (2006): «Reflejo espacial del cambio léxico: los atlas lingüísticos y el DCECH», en *Actes del VII Congrès de Lingüística General (Barcelona, 18-21 de abril de 2006)*, publicado en formato digital en CD-ROM.
- ROSALES SOLÍS, M.ª Auxiliadora (2016): «La derivación nominal de los sufijos con valor frecuentativo en el español de América Central», *Revista de Lengua y Literatura*, <http://revistadelenguayliteratura.com/la-derivacion-nominal-de-los-sufijos-con-valor-frecuentativo-en-el-espanol-de-america-central/> [Consulta: 15/06/2020].
- RUIZ TINOCO, Antonio (2018): «Geocorpus del español de las redes sociales y cartografía automática», *Monográficos SINOELE. IX Congreso Internacional de la Asociación de Hispanistas* (Bangkok 2016), 17, pp. 598-608. http://www.sinoele.org/images/Revista/17/monograficos/AAH_2016/AAH_2016_antonio_ruiz.pdf [Consulta: 15/06/2020].
- RUSINÉS GRAMUNT, Ernest (1999): «Crítica internacional de l'obra de Joan Coromines», en Joan Solà (ed.): *L'obra de Joan Coromines. Cicle d'estudi i homenatge*, Sabadell: Fundació Caixa de Sabadell-Aula de Ciència i Cultura, pp. 169-181.
- SALVADOR, Gregorio (1965): «Estudio del campo semántico «Arar» en Andalucía», *Archivum: Revista de la Facultad de Filología*, 15, pp. 73-111.

- SALVADOR, Gregorio (1980): «Lexicografía y geografía lingüística», *Revista Española de Lingüística*, X, pp. 49-58.
- SANCHÍS GUARNER, Manuel (1953): *La cartografía lingüística en la actualidad y el Atlas de la Península Ibérica*. Madrid: CSIC/Instituto Miguel de Cervantes.
- SARALEGUI, Carmen (2003): «La lematización como problema del paso del atlas lingüístico al diccionario», en Carmen Saralegui y Manuel Casado Velarde (coords.): *Pulchre, bene, recte. Estudios en homenaje a Fernando González Ollé*. Navarra: EUNSA, pp. 1311-1331.
- SCOSYA = Smith, Jennifer, David Adger, Brian Aitken, Caroline Heycock, E. Jamieson & Gary Thoms. 2019. *The Scots Syntax Atlas*. University of Glasgow. <https://scotssyntaxatlas.ac.uk> [Consulta: 15/06/2020].
- SOUSA, Xulio (2017): «From field notebooks to automatic mapping: the *Atlas Lingüístico Galego* database», *Dialectologia et Geolingüística*, 23/1, pp. 1-22.
- SOUSA, Xulio (2020): «From Regional Dialects to the Standard: Measuring Linguistic Distance in Galician Varieties», *Languages*, 5/4, pp. 1-14.
- SOUSA, Xulio (en prensa): «Humanidades digitales y geografía lingüística: la edición digital del *Atlas Lingüístico de la Península Ibérica*» en Ángel Gallego y Francesc Roca (eds.): *Dialectología digital*. Anexo de *Verba*.
- TAGLIAVINI, Carlo (1949): «Di alcuni denominazioni della <pupila> (studio di onomasiologia, con speciale riguardo alle lingue camito-semitiche e negro africane)», *Annali dell'Istituto Universitario Orientale di Napoli*, III, pp. 341-378.
- THIBAUT, André (2007): «Otras orientaciones de la lingüística románica a lo largo del siglo XIX», en José Enrique Gargallo y M.^a Reina Bastardas (eds.), *Manual de lingüística románica*, Barcelona: Ariel, pp. 351-373.
- THOMAS, Alan R. (1980): *Areal Analysis of Dialect Data by Computer: A Welsh Example*. Cardiff: University of Wales Press.
- TISATO, Graziano (2019): «Acquisizione Digitale dell'Intero AIS», documento digital. <https://www.aisv.it/aisv2019/abstracts/11.pdf> [Consulta: 15/06/2020].
- TUSÓN, Jesús (2003): *Introducción al lenguaje*, Barcelona: UOC.
- URITANI, Nozomu y Aurora BERRUETA DE URITANI (1985): «Los diminutivos en los atlas lingüísticos españoles», *Lingüística Española Actual*, VII/2, pp. 203-236.
- VENY, Joan (1991): «Petges aragoneses en els dialectes catalans meridionals», en *Mots d'ahir i mots d'avui*. Barcelona: Empúries, pp. 97-114.
- VENY, Joan (2000): «De la nineta a l'ànima de l'ull», en *Jornades de la secció filològica de l'Institut d'Estudis Catalans a Elx i a la Universitat d'Alacant (16 i 17 d'octubre de 1998)*. Institut d'Estudis Catalans/Ajuntament d'Elx, Barcelona/Elx, pp. 83-92.
- VIERECK, Wolfgang (2000): «La historia cultural europea: algunas observaciones en el *Atlas Linguarum Europae*», *Revista Española de Lingüística*, 30/2, pp. 487-503.
- WALTERS, Keith (1990-1992): «Dialectología», en Frederick J. Newmeyer (comp.), *Panorama de la lingüística moderna de la Universidad de Cambridge*. Madrid: Visor, vol. 4, pp. 149- 172.
- WENKER, Georg (1881): *Sprachatlas von Nord-und Mitteldeutschland*. Auf Grund von systematisch mit Hilfe der Volksschullehrer gesammeltem Material aus circa 30 000 Orten. Straßburg.