

DISTANCIA Y PROXIMIDAD PROSÓDICA
ENTRE ALGUNAS VARIEDADES DEL ESPAÑOL:
UN ESTUDIO DIALECTOMÉTRICO A PARTIR DE
DATOS ACÚSTICOS

DISTANCE AND PROSODIC CLOSENESS BETWEEN SOME
VARIETIES OF SPANISH: A DIALECTOMETRIC STUDY FROM
ACOUSTIC DATA

ANA MARÍA FERNÁNDEZ PLANAS

Lingüística General. Universitat de Barcelona. Barcelona, España
anamariafernandez@ub.edu

JOSEFA DORTA

Lingüística General. Universidad de La Laguna. Tenerife, España
jdorta@ull.edu.es

PAOLO ROSEANO

Lingüística General. Universitat de Barcelona; Universitat Pompeu Fabra.
Barcelona, España
paolo_r@hotmail.it

CHAXIRAXI DÍAZ

Lingüística General. Universidad de La Laguna. Tenerife, España
chadiaz@ull.edu.es

WENDY ELVIRA-GARCÍA

Lingüística General. Universitat de Barcelona. Barcelona, España
wendyelviragarcia@gmail.com

JOSÉ ANTONIO MARTÍN GÓMEZ

Lingüística general. Universidad de La Laguna. Tenerife, España
josemartgo@gmail.com

EUGENIO MARTÍNEZ CELDRÁN

Lingüística General. Universitat de Barcelona. Barcelona, España
martinezceldran@ub.edu

RESUMEN

Este trabajo somete los datos de F0 de oraciones declarativas e interrogativas del corpus experimental de AMPER, emitidas por sujetos de seis variedades del español peninsular y siete del español insular, a un análisis dialectométrico con dos rutinas diferentes. Por un lado, el método denominado Calcu-Dista, del Laboratori de Fonètica de la Universitat de Barcelona, trabaja con una matriz de distancias de datos cuantitativos de F0 en semitonos; para representar las distancias prosódicas sin necesidad de recurrir a esa gran matriz de datos se utilizan dendrogramas creados en SPSS a partir del análisis de clúster o de agrupación y gráficos del tipo MDS (Multi Dimensional Scaling), que representan las distancias entre los puntos de encuesta en un espacio virtual. Por otro lado, la herramienta en línea de fuente abierta DiaTech, desarrollada por el equipo de investigación de la UPV/EHU, permite realizar análisis de distancias con datos de tipo nominal, por lo que se ha usado el sistema métrico autosegmental para el etiquetaje prosódico de las curvas de F0; también representa los resultados mediante gráficos como, por ejemplo, el dendrograma. El análisis de los mismos datos con ambos métodos, cuantitativo y cualitativo, permite observar la distribución de las diferentes variedades estudiadas y, además, consolida los resultados obtenidos ya que, con algunas diferencias, ambos coinciden en separar los puntos de encuesta en dos grandes grupos bastante alejados entre sí tanto para la modalidad interrogativa como para el conjunto de las dos modalidades; en las declarativas, en cambio, no se han formado grupos tan coherentes.

Palabras clave: AMPER, análisis dialectométricos, entonación, cluster.

ABSTRACT

This paper presents the results of an dialectometric analysis of F0 data of interrogative and declarative sentences from AMPER experimental corpus, uttered by speakers of six varieties of peninsular Spanish and seven speakers of insular Spanish, according to two different routines. The Calcu-Dista method, developed by the Laboratori de Fonètica of the University of Barcelona, works with a distance matrix of quantitative F0 data in semitones. To represent prosodic distances without the need to use such large data matrix, we used SPSS dendrograms created through cluster analysis and MDS (Multi Dimensional Scaling) type graphics, which represent the distances between sample points in a virtual space. On the other hand, the open source online tool DiaTech, developed by the UPV / EHU research team, allows to analyze distances using nominal data, though. To solve this problem, we have used the autosegmental metric system to label F0 curves. This software also represents results using dendrograms. The analysis of the available data using these two different quantitative and qualitative methods allows to represent graphically the distribution of the different studied varieties of Spanish. Furthermore, both methods are capable of sort out the available data into two distinct groups, for both the interrogatives and interrogative-declarative sets; while in the case of the declarative sentences, no coherent groups have been formed.

Keywords: AMPER, dialectometric analysis, intonation, cluster.

Recibido: 22.05.2015. *Aceptado:* 09.09.2015.

1. INTRODUCCIÓN

1.1. La dialectometría

En la moderna dialectología, disciplina clasificatoria de carácter instrumental de datos geográficos lingüísticos a partir de procedimientos objetivos estadísticos, cada vez más desde los años 70 del siglo XX se impone la técnica dialectométrica (Séguy, 1971) como herramienta para tratar los datos. Goebel (1981) define la dialectometría como una alianza metodológica entre la geolingüística y la taxonomía numérica como disciplina matemática. Exactamente el autor lo expone de forma sintética de la siguiente manera: dialectometría = geografía lingüística + taxonomía numérica (Goebel, 1981: 349). Ciertamente, lo que los estudios dialectométricos pretenden es utilizar una enorme cantidad de datos que se han generado a través de los estudios dialectológicos y los atlas lingüísticos para *establecer agrupaciones entre la masa de datos empíricos* disponibles y obtener una *distribución del espacio virtual de los datos* (Fernández Planas, Roseano, Martínez Celdrán y Romera Barrios, 2011: 145). En definitiva, el objetivo fundamental consiste en realizar un tipo de análisis estadístico para el cálculo de las distancias dialectales entre distintas lenguas o variedades de una misma lengua, para lo cual la dialectometría funciona operando con cantidades grandes de datos de forma cuantitativa y objetiva. Una ventaja de este tipo de análisis reside en la plasmación gráfica de los resultados, normalmente en forma de árbol invertido (dendrograma) pero, también, en un mapa virtual de distancias obtenido a partir de escalamiento multidimensional (o a partir de otros procedimientos estadísticos), que permite una rápida asociación entre los elementos considerados a partir de su cercanía o su lejanía —es decir, de sus semejanzas o sus diferencias— y posibilita condensar una gran cantidad de información cuantitativa en un espacio relativamente reducido. A pesar de sus ventajas para tratar cantidades ingentes de datos de forma comprensible y fácilmente aprehensible y para estudiar los datos sin estar influidos por apriorismos, se puede objetar a la técnica el no tener en cuenta diferencias cualitativas entre los ellos (Clua, 1999), es decir, no observar que ciertas diferencias lingüísticas son más relevantes que otras cualitativamente, igual que puede ser discutible la medida de similitud elegida para computar las distancias. A pesar de trabajarse la técnica desde hace varias décadas, todavía no ha sido aceptada ni comprendida por todos los dialectólogos, seguramente por el hecho de obviar las diferencias cualitativas entre las variables sometidas a estudio. Sin embargo, la dialectometría no pretende eliminar el estudio dialectológico tradicional sino que busca completarlo y erigirse como una herramienta esencialmente útil cuando se manejan cantidades enormes de datos.

Por lo que respecta a las lenguas romances, el método se ha aplicado principalmente a las áreas lingüísticas del ladino (Goebel, 1993; Bauer, 2005), el italiano

(Bauer, 2003), el francés (Séguy, 1973; Verlinde, 1988; Goebel, 1987, 2003), el gallego (Álvarez Blanco, Dubert y Sousa, 2006; Sousa, 2006; Saramago, 2002), el bable (D'Andrés, Álvarez-Balbuena y Suárez Fernández, 2007) o el catalán (Clua, 2004; Polanco, 1992). Fuera de la Romania se utiliza también en estudios dialectológicos de lenguas como el holandés (Heeringa y Nerbonne, 2001), el inglés (Goebel y Schiltz, 1997) o el euskara (Aurrekoetxea, 1992).

Como se ve, los trabajos y los grupos de estudio implicados son relativamente abundantes, pero ninguno de ellos se dedica especialmente a aspectos prosódicos de las variedades que analiza (aunque a veces sí se fundamentan en datos fonéticos segmentales), a diferencia del trabajo que se presenta a continuación.

1.2. El Atlas Multimedia de Prosodia del Espacio Románico (AMPER)

El *Atlas Multimedia de Prosodia del Espacio Románico* (AMPER) es un marco adecuado para el sometimiento de los datos que lo sustentan a un estudio dialectométrico. En efecto, este macroproyecto, impulsado desde el *Centre de Dialectologie de l'Université Stendhal Grenoble 3* por Michel Contini, Antonio Romano y Jean Pierre Lai (Contini, 1992; Contini, Lai, Romano y Rouillet, 1998; Contini et al., 2002; Romano, Lai y Rouillet, 2005; Romano y Contini, 2001; Contini, Lai y Romano, 2002; Romano, 2003; Fernández Planas, 2005), tiene por objetivo el análisis prosódico de todas las lenguas y variedades románicas habladas en Europa y América con el propósito de que los datos obtenidos sirvan, en primera instancia, para sustentar el atlas proyectado en el entorno multimedia (<http://w3.u-grenoble3.fr/dialecto/AMPER/amper.htm>). Además, se pretende también poder realizar comparaciones amplias y detalladas sobre las lenguas y variedades estudiadas, objetivo este último para el que la dialectometría se vislumbra como una herramienta muy eficaz, ahora que en algunos dominios lingüísticos el proyecto está bastante avanzado, dado que permite la agrupación y distribución de los datos obtenidos indicando su lejanía o proximidad.

En el marco global de AMPER, el español y sus variedades está siendo ampliamente analizado por diversos grupos de investigación coordinados de manera general por Eugenio Martínez Celdrán (Laboratori de Fonètica de la Universitat de Barcelona), con la subcoordinación de Josefa Dorta (Laboratorio de Fonética de la Universidad de La Laguna) y de Yolanda Congosto (Universidad de Sevilla). La amplitud de territorios en los que se habla esta lengua y el número de grupos de investigación y de investigadores implicados han permitido obtener una gran cantidad de datos prosódicos gráficos y numéricos de frecuencia fundamental, duración e intensidad, por lo que es posible iniciar, entre otros, estudios dialectométricos diversos relacionados con las modalidades enunciativa neutra e interrogativa absoluta o total que son, por el momento, las únicas estudiadas en AMPER.

2. OBJETIVOS E HIPÓTESIS

Este trabajo persigue dos objetivos, uno de carácter metodológico y el otro, teórico-práctico. Por una parte, desde el punto de vista metodológico se pretende presentar a la comunidad fonética internacional una herramienta para llevar a cabo análisis dialectométricos con datos prosódicos cuantitativos. Se trata de una herramienta llamada Calcu-Dista y desarrollada en el Laboratori de Fonètica de la Universitat de Barcelona. Por otra parte, el objetivo teórico-práctico de este trabajo consiste en someter a un estudio dialectométrico los datos de F0 de oraciones emitidas en las dos modalidades ya referidas por sujetos de algunas de las variedades del español peninsular (Martínez Celdrán y Fernández Planas, Eds., 2003-2015) e insular (Dorta, Ed., 2003-2015; Dorta, 2009) con el propósito de que el análisis permita establecer sus relaciones de proximidad o distancia prosódica y postular una explicación diacrónica para estas relaciones. Los resultados teórico-prácticos, obtenidos tanto con la herramienta metodológica que se presenta (Calcu-Dista) como con otra herramienta novedosa pero ya conocida y aceptada en el mundo de la dialectometría –DiaTech–, permitirán seguir avanzando en esta línea metodológica, por lo que en futuros trabajos será posible ampliar el número de datos respecto de los aquí considerados.

De acuerdo con los dos objetivos, las hipótesis de partida del presente trabajo son también dos. Por un lado, la hipótesis metodológica postula que Calcu-Dista permitirá clasificar geolingüísticamente los datos a partir de sus similitudes, como hacen otros sistemas (probablemente uno de los más conocidos es el método de Hans Goebel desarrollado en la Universidad de Salzburgo, pero existen otras aproximaciones metodológicas como la de la escuela holandesa de la Universidad de Gröningen). La novedad de este sistema frente a los demás reside en que funciona a partir de datos numéricos de parámetros prosódicos (entre los parámetros prosódicos fundamentales –frecuencia fundamental (F0), duración e intensidad– se centra en el del movimiento entre dos pausas de la frecuencia fundamental, conocido como entonación), lo que lo convierte en un método óptimo para estudios llevados a cabo en el marco del AMPER. Otro punto a favor de dicho sistema reside en que para trabajar este método aprovecha herramientas al alcance de cualquier fonetista con un nivel de usuario informático (Praat, Excel y SPSS) con un protocolo claramente establecido y unas rutinas de automatización de procesos creadas *ad hoc*.

Por otro lado, la hipótesis teórico-práctica que se plantea en el presente trabajo postula que la modalidad interrogativa será más decisiva que la enunciativa a la hora de clasificar los diferentes puntos de encuesta (como ha sucedido en trabajos anteriores para el catalán –por ejemplo Fernández Planas et al., 2011–, o para el español de España y de América –por ejemplo, Dorta, Ed., 2013– y como parece ser desde el punto de vista perceptivo –por ejemplo, Fernández Planas et al., en

prensa-) y que estos se distribuirán en dos grandes grupos, uno de ellos integrado fundamentalmente por los puntos peninsulares y balear, y otro por el punto de encuesta andaluz y por el resto de zonas insulares por razones históricas relacionadas con la colonización de Canarias y el descubrimiento de América.

3. METODOLOGÍA

3.1. Los puntos de encuesta

Para llevar adelante este estudio se han elegido catorce puntos de encuesta que se clasifican de diversa manera entre ellos. Por su distribución geográfica, seis corresponden al español peninsular (Barcelona, Lleida, Palencia, Salamanca, Madrid y Bullas) y siete al español insular (Palma de Mallorca, Santa Cruz de La Palma, San Cristóbal de La Laguna, Las Palmas de Gran Canaria, Arrecife, La Habana y Santiago de Cuba). Por las lenguas que se hablan en dichas zonas, tres corresponden a zonas bilingües catalán-español que se relacionan con tres variedades muy bien delimitadas de catalán: catalán oriental (Barcelona), occidental (Lleida) y balear (Palma), y diez a zonas monolingües de español (Palencia, Salamanca, Madrid, Bullas, Santa Cruz de La Palma, San Cristóbal de La Laguna, Las Palmas de Gran Canaria, Arrecife, La Habana y Santiago de Cuba). Entre estas últimas, tres corresponden a puntos del centro peninsular (Palencia, Salamanca y Madrid), dos a la zona meridional peninsular (Bullas –Murcia– y Granada), cuatro se sitúan en el archipiélago canario y dos en la isla de Cuba, en el Caribe. Entre los puntos canarios, dos corresponden a la provincia oriental, uno en cada una de las islas mayores (Lanzarote y Gran Canaria) y dos a la provincia occidental, uno en cada una de las islas mayores (Tenerife y La Palma). Entre los dos de Cuba, uno representa la zona occidental de la isla (La Habana) y el otro, la zona oriental (Santiago de Cuba). La localización sobre el mapa de estos puntos aparece en las figuras 1, 2 y 3.



Figura 1. Localización de los puntos de encuesta estudiados en la península y las Islas Baleares.

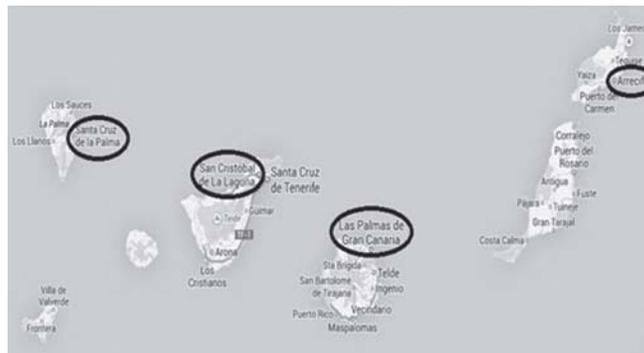


Figura 2. Localización de los puntos de encuesta estudiados en las Islas Canarias.



Figura 3. Localización de los puntos de encuesta estudiados en la isla de Cuba en el Caribe americano.

3.2. Los sujetos

El estudio se realiza sobre voces masculinas y femeninas. En la mayoría de los puntos de encuesta se cuenta con un representante de cada sexo, pero en algunos solamente se dispone de la voz femenina (Santa Cruz de La Palma, Palencia) o de la voz masculina (Salamanca). La Tabla I muestra los códigos AMPER¹ de los distintos puntos de encuesta y sujetos utilizados en los análisis de este trabajo.

Tabla I. Códigos de las variedades de español utilizadas, sujetos y puntos de encuesta.

Código	Variedad de español	Inform.	Punto de encuesta
wc92	Español canario occidental	H	San Cristóbal de La Laguna (Tenerife)
wc91	Español canario occidental	M	San Cristóbal de La Laguna (Tenerife)
wcm1	Español canario occidental	M	Santa Cruz de La Palma (La Palma)
wck2	Español canario oriental	H	Las Palmas de Gran Canaria
wck1	Español canario oriental	M	Las Palmas de Gran Canaria
wcp2	Español canario oriental	H	Arrecife (Lanzarote)
wcp1	Español canario oriental	M	Arrecife (Lanzarote)
waqh2	Español caribeño cubano	H	La Habana (Cuba)
waqh1	Español caribeño cubano	M	La Habana (Cuba)
waqs2	Español caribeño cubano	H	Santiago de Cuba
waqs1	Español caribeño cubano	M	Santiago de Cuba
wp02	Español de Catalunya	H	Barcelona (Catalunya)
wp01	Español de Catalunya	M	Barcelona (Catalunya)
wp12	Español de Catalunya	H	Lleida (Catalunya)
wp11	Español de Catalunya	M	Lleida (Catalunya)
wx02	Español balear	H	Palma (Mallorca)
wx01	Español balear	M	Palma (Mallorca)
wu12	Español murciano	H	Bullas (Murcia)
wu11	Español murciano	M	Bullas (Murcia)
wm02	Español madrileño	H	Madrid
wm01	Español madrileño	M	Madrid
wl01	Español de Castilla y León	M	Palencia (Castilla y León)
wl12	Español de Castilla y León	H	Salamanca (Castilla y León)

¹ Estos códigos no son necesarios para la correcta lectura del trabajo pero son útiles en la comprobación visual y perceptiva que se puede hacer de estos puntos en el *Atlas Multimedia de Prosodia del Espacio Románico* (véanse: Dorta (Ed.) (2003-2015) y también Martínez Celdrán y Fernández Planas (Eds.) (2003-2015)).

Todos los sujetos que han participado en este estudio cumplen con los requisitos etarios, dialectales y culturales establecidos en el proyecto AMPER-en-España-e-Iberoamérica. Es decir, todos ellos tienen entre 25 y 55 años, carecen de estudios superiores, son nativos de la zona que representan e hijos de personas con el mismo perfil.

3.3. El corpus

El corpus utilizado consiste en 54 frases por sujeto del llamado corpus fijo en el marco de los estudios realizados para el *Atlas Multimedia de Prosodia del Espacio Románico (AMPER)*, cercano al habla de laboratorio, conseguido mediante elicitación textual. Se trata de las frases que en dicho proyecto se conocen como “sin expansión”, es decir, constan de tres acentos léxicos distribuidos en un sujeto, un verbo y un complemento verbal. Las posiciones de sujeto y complemento contemplan todas las posibilidades acentuales (palabras oxítonas, paroxítonas y proparoxítonas) en todas las combinaciones posibles. La posición del verbo siempre es paroxítona. Cada frase se repitió tres veces, con lo cual las nueve posibilidades combinatorias por modalidad (enunciativa e interrogativa absoluta) se emitieron tres veces por sujeto. Ejemplos de estas frases con la misma estructura sintáctica son *La guitarra se toca con paciencia*, *El saxofón se toca con obsesión* y *La cítara se toca con pánico*. Todas ellas han sido producidas de manera pragmática y emocionalmente neutra, tal como está establecido en el marco del macroproyecto internacional AMPER.

En total, el corpus de este trabajo ha quedado constituido por 1.242 oraciones, emitidas de la forma más natural posible en un tempo cómodo para los sujetos y sin hacer pausas en el interior de las frases. De cada una de las frases se han trabajado tres valores de F0 por vocal, así es que el número de datos final ha sido de más de 40.000².

La equivalencia del corpus en los diferentes puntos de encuesta, incluso entre lenguas diferentes, en el seno de AMPER respecto a la estructura sintáctica, a la variación y la combinatoriedad acentual, a las modalidades estudiadas y a la consideración pragmática de las emisiones, hacen viable la comparación prosódica entre ellas.

² En el marco del macroproyecto AMPER, como se ha dicho, el análisis de cada una de las repeticiones de las frases, así como de las medias entre las diferentes repeticiones del mismo enunciado, contempla el estudio de los parámetros F0 (con tres valores por vocal), duración e intensidad (con un valor en cada una de las variables por vocal). En este trabajo, sin embargo, nos ceñimos exclusivamente al estudio del tono fundamental.

3.4. Los programas para el análisis dialectométrico utilizados

La dialectología ha trabajado tradicionalmente de forma mayoritaria con datos morfológicos o léxicos y, desde el punto de vista fonético-fonológico, con datos segmentales, normalmente alfabéticos y no numéricos. Ha necesitado, por tanto, algoritmos capaces de proporcionar una medida cuantitativa de las diferencias-similitudes entre dialectos a partir de variables nominales. Sin embargo, los estudios entonativos acústicos, enfocados desde un punto de vista eminentemente fonético, operan con variables numéricas (el parámetro F0, que interesa en este trabajo, se suele expresar en Hz o en alguna medida normalizadora, como los semitonos) y, por ello, puede contar con las medidas normales de distancias entre datos de partida ya numéricos.

En el presente trabajo se han analizado los datos de dos formas: en primer lugar, con el método que se ha denominado Calcu-Dista (Roseano, Elvira-García, Fernández Planas y Martínez Celdrán, en preparación) y, en segundo lugar, con la herramienta en red DiaTech (véanse los apartados 3.4.2 y 3.4.3.). En el primero, Calcu-Dista, se trabajó con los datos cuantitativos numéricos del parámetro de F0 en semitonos³. En el segundo, DiaTech, una herramienta novedosa ya validada y aceptada en círculos dialectológicos internacionales, se trabajó de forma cualitativa a partir de los etiquetajes entonativos de las mismas melodías obtenidos en el nivel superficial de su análisis (véase al respecto el apartado 3.4.4.) de las curvas melódicas. La consideración de los datos de dos formas distintas, consolida los resultados obtenidos puesto que se comprueba que, en lo fundamental, van en la misma dirección. Este hecho valida tanto el objetivo metodológico propuesto como el teórico respecto a la clasificación de los datos prosódicos sometidos a análisis que, a su vez, constituye una práctica del objetivo metodológico.

3.4.1. Preparación de los archivos

Antes de proceder a trabajar con los datos disponibles tanto en Calcu-Dista como en DiaTech ha sido necesario preparar los archivos txt de partida, obtenidos en el análisis de los datos mediante el programa AMPER-2006 (López Bobo et al., 2007) en el entorno Matlab, basado en unas primeras rutinas informáticas (AMPER-dat y AMPER-fox) creadas *ad hoc* para analizar la prosodia de las lenguas

³ En el seno del proyecto AMPER, existe el programa Stat-Distances, creado por A. Rilliard *ad hoc* en el entorno MatLab para calcular las distancias objetivas entre los datos acústicos prosódicos de los diversos puntos de encuesta (Moutinho et al., 2011; Romano y Miotti, 2008; Rilliard y Lai, 2008; Romano, Contini, Lai y Rilliard, 2011). Tiene un enfoque netamente fonético y calcula las distancias interdialectales a partir de una medida de correlación basada en múltiples valores de F0 (tres para cada vocal) en cada frase. Este método de cálculo de distancias entre parámetros acústicos parte de Hermes (1998a, 1998b), quien introduce una ponderación en función de la intensidad. El programa está inspirado en la idea inicial propuesta en la tesis de A. Romano (1999).

románicas (Romano e Interlandi, 2005).

En el caso del análisis con DiaTech los datos, además, se han pasado por tres programas creados en el Laboratori de Fonètica de la UB: AMPERReno, AMPERExtra y AMPEReti (Roseano, 2012). El primero de ellos permite automáticamente cambiar el nombre de los archivos de tal forma que el nuevo nombre informa no solo de la localidad, el sujeto, la repetición de la frase, la modalidad oracional y su estructura sintáctica, sino también del tipo acentual de cada acento léxico, del número total de sílabas de la oración, de la posición de las sílabas tónicas en ella y del timbre de dichas vocales tónicas. El segundo, AMPERExtra, permite extraer a un archivo Excel todos los datos en forma de matriz. El tercer programa etiqueta automáticamente la prosodia de las oraciones en el marco métrico autosegmental, a partir de la consideración de un conjunto de fórmulas basadas en una primera versión de los algoritmos de etiquetaje automático entonativo (Martínez Celadrán y Fernández Planas, 2003). Puede encontrarse una explicación más detallada del funcionamiento de estos tres programas en Roseano (2012) y Roseano, Fernández Planas y Martínez Celadrán (2014).

3.4.2. *Calcu-Dista*

Esta metodología de análisis consiste en una rutina para el cálculo de distancias prosódicas a partir de datos numéricos de las melodías estudiadas en semitonos propuesta por el *Laboratori de Fonètica* de la UB (Roseano et al., 2014) e incluye el trabajo con tres programas bien conocidos: Praat v. 5.4.01 (Boersma y Weenink, 2014), Excel (Microsoft Office 2007) y SPSS Statistics 20 (IBM), tras el paso por tres programas creados en el seno del *Laboratori de Fonètica* de la UB y circunscritos al ámbito AMPER: AMPERReno, AMPERExtra y AMPEReti (Roseano, 2012).

En primer lugar, un *script* de Praat creado *ad hoc* extrae, a partir de los archivos txt de cada repetición de las frases proporcionados por AMPER-2006, los valores de F0 en semitonos y los coloca en una matriz de datos comparando cada repetición de una frase en un mismo hablante y entre hablantes distintos considerando tres valores por vocal.

En segundo lugar, un análisis en Excel sobre la salida de Praat aplica la fórmula de las distancias escogida (véase la Figura 4). En este trabajo, de acuerdo con la propuesta clásica de Hermes (1998a), se escogió como índice de la distancia entonativa entre dos frases –para el caso x e y –, la media cuadrática de la diferencia entre los valores de F0 de la frase x y de la frase y en cada uno de los puntos de medición. Para los dos conjuntos x e y de valores de F0 $\{f0_{x_1}, f0_{x_2}, \dots, f0_{x_N}\}$ y $\{f0_{y_1}, f0_{y_2}, \dots, f0_{y_N}\}$, donde N es el número de puntos de medición de F0 en cada una de las dos frases, mientras que $f0_{x_i}$ y $f0_{y_i}$ son los valores de F0 en st en cada uno de esos puntos. La media cuadrática en cuestión viene dada por la fórmula que aparece en la Figura 4.

$$RMS = \sqrt{\frac{\sum_{i=1}^N (f0x_i - f0y_i)^2}{N}}$$

Figura 4. Fórmula empleada para calcular las distancias.

Esta fórmula proporciona la distancia entre dos frases con la misma estructura (por ejemplo, entre dos declarativas SVO con sujeto llano, verbo llano y objeto esdrújulo) de dos puntos de encuesta.

Para determinar la distancia general entre todas las frases de dos puntos de encuesta, puesto que la distribución de las distancias no es normal, de acuerdo con Moutinho *et al.* (2011: 44) se ha escogido la mediana de las RMS calculadas por cada pareja de frases x e y . A partir de las medianas de las distancias entre cada par de puntos de encuesta se puede construir una matriz de distancias como la que aparece en la Tabla III de los resultados.

En tercer lugar, la matriz de distancias constituye a su vez la base para la fase final del proceso de análisis, que se efectúa con SPSS y consiste en un análisis de conglomerados *cluster*, técnica multivariante cuya finalidad es clasificar los puntos de encuesta en grupos a partir de la semejanza entre sus características entonativas donde como método de comparación se utiliza la media de las distancias entre los grupos tratando de lograr la máxima homogeneidad en cada grupo y la mayor diferencia entre los grupos. Como medida utilizamos el intervalo de distancia euclidiana.

En Calcu-Dista el análisis de *cluster* se ha realizado usando como variables los datos de distancias creados previamente y los casos se han etiquetado mediante la variable de nombre de población. A partir de la matriz de datos creada, se han calculado las diferencias mediante una medida de disimilitud o de distancia habitual, esto es, la *distancia euclidiana* según la fórmula siguiente:

$$d(x,y) = \|x-y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Figura 5. Fórmula de la función distancia euclídea.

Por tanto, se presupone que los grupos de *cluster* resultantes estarán formados por variedades parecidas, es decir, que la distancia entre los dialectos en un mismo grupo será pequeña. Para clasificar los elementos en *clusters* se ha utilizado el método acumulativo que permite formar grupos haciendo *clusters* cada vez más grandes, es decir, que se parte de los individuos para su agrupación. Y como se dijo

más arriba, el método de conglomeración usado (intergrupos) calcula la distancia entre dos grupos a partir de la media aritmética de las distancias posibles entre esos grupos.

En SPSS se obtienen también los gráficos deseados. Particularmente, interesan dos tipos de ellos generados a partir de los cálculos matemáticos sobre los resultados cuantitativos numéricos: un dendrograma, creado a partir del análisis de *cluster*, y un gráfico de escalamiento multidimensional (EMD o MDS –*Multi-Dimensional Scalling*–). El dendrograma permite ver, en forma de árbol invertido, cómo se agrupan los datos hasta el nivel que se considera oportuno. El MDS representa bidimensionalmente de forma gráfica las distancias entre los sujetos o puntos de encuesta de la manera más objetiva posible en un espacio virtual. Este método estadístico pretende construir un espacio métrico con el menor número de dimensiones posibles, de tal manera que permite representar las proximidades o preferencias entre los objetos con el mayor grado de fidelidad. Desde un conjunto de objetos se establecen sus propiedades numéricas a partir de las cuales se elaboran las tablas de proximidad (o de similitud) y, finalmente, se trasladan estas proximidades a un espacio, un mapa de objetos (Matas, 2006). En realidad, ambos tipos de gráficos –dendrogramas y espacios MDS– proporcionan la misma información, y así se puede comprobar en el apartado de resultados. La ventaja de ambas formas de representación es la de permitir captar la distribución y la agrupación de los datos sin necesidad de tener que recurrir a una matriz de distancias numérica de proporciones enormes.

3.4.3. *DiaTech*

La herramienta en línea de fuente abierta *DiaTech* (<http://eudia.ehu.es/diatech/index/>), desarrollada por el equipo interdisciplinar de investigación de la UPV/EHU encabezado por G. Aurrekoetxea, parte de una base de datos relacional, usa diversos análisis estadísticos y permite construir diferentes tipos de mapas en los que visualizar las diferencias lingüísticas entre diferentes áreas o puntos dialectales. Incorpora la posibilidad de utilizar múltiples tipos de análisis (*cluster* o correlación, por ejemplo) y diferentes algoritmos de clasificación y formas de visualización de los resultados del análisis estadístico. Desde su reciente aparición, se está constituyendo en una herramienta muy valorada por los dialectólogos. Sus características pueden encontrarse en Aurrekoetxea, Fernández-Aguirre, Rubio, Ruiz y Sánchez (2013).

Como ya se dijo en la introducción, el análisis dialectométrico en el marco de la dialectología estricta se suele efectuar a partir de datos morfológicos o léxicos (también segmentales en el nivel fónico), es decir, alfabéticos. Pero los datos entonativos acústicos se obtienen de forma numérica (sea en Hz o en st). Por ello, para analizar datos entonativos con las técnicas dialectométricas comúnmente

utilizadas, o con la reciente herramienta DiaTech, es preciso convertir los datos numéricos en cadenas alfabéticas. Con este fin se etiquetaron las melodías objeto de interés en cadenas nominales según los postulados de la teoría métrico-autosegmental. Dentro del modelo entonativo métrico autosegmental (Pierrehumbert, 1980), que se siguió en este trabajo, se suele transcribir la entonación con una serie de símbolos alfabéticos que son combinaciones de tonos altos (H, del inglés *High*) y bajos (L, del inglés *Low*) en los acentos léxicos prominentes o destacados de la frase, donde el asterisco después del símbolo (*) señala la sílaba tónica y donde el símbolo del porcentaje (%) y el guión (–) se refieren a los tonos de frontera o de juntura, respectivamente frontera inicial o final y frontera intermedia.

A partir de estas cadenas nominales, adaptadas a los requisitos de funcionamiento de DiaTech, se trabajaron las distancias entre los puntos de encuesta. En el presente estudio se ha utilizado un análisis de *cluster*, el índice de distancia lingüística de Levenshtein y el algoritmo de distancia estadística Ward.

3.4.4. *El etiquetaje métrico autosegmental*

El estudio dialectométrico en DiaTech se ha realizado, como se ha dicho, a partir de las etiquetas entonativas resultantes de la interpretación de las melodías en un nivel superficial, es decir, plenamente fonético y no subyacente o profundo, más cercano a la fonología, que han sido transformadas en *strings* alfabéticas que se pueden analizar como variables nominales⁴. Dentro del marco teórico métrico autosegmental (AM), existen propuestas de sistemas de notación (ToBI o Tones and Breaks Indices, Beckman y Hirschberg, 1994) para diferentes lenguas. El sistema para la transcripción de la entonación del español, Sp_ToBI (Estebas-Vilaplana y Prieto, 2009; Prieto y Roseano, 2010) se suele utilizar para efectuar una notación fonológica pero también, aunque menos frecuentemente, una notación fonética, como la que se utiliza en este trabajo (basada en Fernández Planas y Martínez Celdrán, 2003) que considera la pertinencia del umbral psicoacústico establecido en 1'5 st (Pamies, Fernández Planas, Martínez Celdrán, Ortega Escandell y Amorós Céspedes, 2001; Rietveld y Gussenhoven, 1985).

Las etiquetas que se utilizan en la transcripción fonética de la entonación utilizando Sp_ToBI en el marco métrico autosegmental tienen una característica formal que las convierte en problemáticas a la hora de aplicar los logaritmos dialectométricos de DiaTech porque tienen un número de caracteres variable entre dos y

⁴ En realidad, pues, estamos estudiando los mismos datos melódicos por un lado en su versión numérica, en bruto, mediante el método Calcu-Dista y, por otro lado, en su interpretación fonética a partir del etiquetaje ToBI en el marco de la teoría autosegmental métrica (Pierrehumbert, 1980; Ladd, 2008). Este etiquetaje se puede abordar en diferentes niveles (Elvira-García et al., 2014; Roseano y Fernández Planas, 2013), el más cercano a los valores numéricos es el puramente fonético pero incluso así, por constituir una interpretación de los mismos, no es exactamente igual que ellos.

nueve (véase la Tabla II). Por esta razón, se ha procedido previamente a normalizar a tres el número de caracteres de cada etiqueta, sin perder ningún detalle entonativo. La Tabla II contiene las etiquetas entonativas fonéticas o de nivel superficial (Fernández Planas y Martínez Celdrán, 2003) y su correspondencia normalizada, tanto para los acentos tonales como para los tonos de juntura. Esta equivalencia normalizada según los requisitos de funcionamiento de DiaTech es la que se ha utilizado efectivamente en el análisis llevado a cabo con esta herramienta.

Tabla II. Etiquetas entonativas en el nivel fonético o superficial (a la izquierda) y normalización de las mismas para su uso en DiaTech (a la derecha).

Etiquetas fonéticas según Sp_TOBI en el marco AM	Etiquetas normalizadas para su uso en DiaTech
L*	OLO
H*	OHO
H*+L	OHL
H+L*	HLO
L*+H	OLH
L+H*	LHO
H+(L*+!H)	HLH
H+(L*+!H)	HLH
(H+L)+H*	MLH
L+(H*+H)	LHM
(L+H*)+H	LHH
H+(L*+L)	HLL
(H+L*)+L	HLL
L+(H*+L)	LHM
(L+H*)+L	MHL
L%	LLL
H%	HHH
%L	LLL
%H	HHH
L-	LLL
H-	HHH

4. RESULTADOS

Como se ha explicado en el apartado de metodología, en este trabajo se ha utilizado tanto Calcu-Dista como DiaTech para el análisis de distancia prosódica entre variedades del español que se hablan en zonas monolingües y bilingües. Los resul-

tados obtenidos con ambos tipos de metodologías muestran algunas diferencias aunque, como se ha adelantado ya (apartado 3.4), coinciden en lo fundamental.

4.1. Calcu-Dista

4.1.1. Enunciativas

La estructura de los conglomerados o *clusters* de las enunciativas de forma jerárquica se puede ver en el dendrograma⁵ del Gráfico 1 obtenido en SPSS.

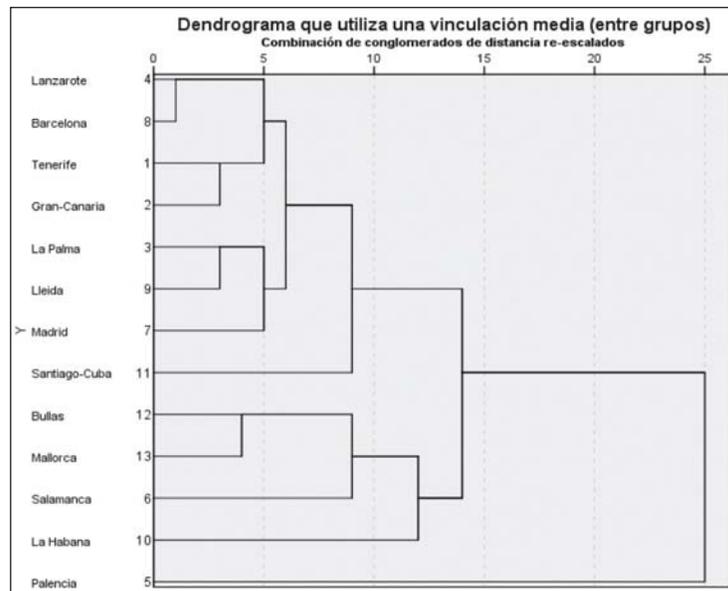


Gráfico 1. Dendrograma de las enunciativas obtenido en SPSS.

⁵ Los dendrogramas constituyen una aplicación directa de un procedimiento clasificatorio habitual en muchas ciencias naturales y humanas. Es el elemento central de la llamada clasificación jerárquica aglomerativa con ramificaciones binarias que se calculan a partir de los datos de la matriz de similitud y por medio de un apropiado algoritmo agrupador de clustering jerárquico (Goebel, 2013). El programa primero calcula las distancias entre cada par de clases en el archivo de entrada. Después, se fusiona iterativamente el par de clases más cercano y se fusiona sucesivamente el siguiente par de clases más cercano y el siguiente más cercano hasta que se fusionan todas las clases. Después de cada fusión, se actualizan las distancias entre todos los pares de clases. Las distancias a las que se fusionan las clases se utilizan para construir un dendrograma. La clase fusionada se produce utilizando el valor medio y la varianza de forma conjunta (ArcGIS, 2014). El análisis de cluster jerárquico puede emplear variables de tipo cuantitativo y de tipo cualitativo, lo cual abre posibilidades para crear estrategias enmarcadas en la lógica del multimétodo y la combinación de lo cualitativo y lo

En este gráfico podemos apreciar que las variedades analizadas se distribuyen en *clusters* cuya formación parece, en principio, incoherente en tanto que las distancias entre las insulares y peninsulares se dan sin aparente justificación. Así, por ejemplo, mientras que Tenerife y Gran Canaria aparecen muy próximas en un mismo *cluster* de manera razonable por pertenecer a las islas Canarias –donde las diferencias son pequeñas, según evidencian los datos precedentes–, Lanzarote y Barcelona también quedan agrupadas en un mismo *cluster* a pesar de que esperaríamos que aparecieran a relativa poca distancia respecto de las otras islas Canarias y de las variedades peninsulares, respectivamente. Los datos de La Palma, Lleida y Madrid también se muestran juntos en un primer nivel. En suma, tienden a formarse dos grupos grandes, uno con más representantes insulares que peninsulares y otro al revés, y otro grupo unimembre para el punto de Palencia.

Se han realizado otras pruebas estadísticas con las distancias desde Granada (punto escogido aleatoriamente) respecto a los demás puntos. En las enunciativas, la prueba de Kolmogorov-Smirnov determina que la muestra no es normal, a partir de aquí la prueba de comparación entre los puntos se hace con test no paramétricos. La prueba de Kruskal-Wallis (Chi-cuadrado=817,480; gl=13; sig=0,000) señala que la diferencia entre todos los grupos es significativa. Por ello son puntos elegidos para formar parte del *Atlas Multimedia de Prosodia del Espacio Románico* (AMPER) que pretende plasmar en un mapa las diferencias acústicas (mostradas gráficamente) y perceptivas entre diferentes variedades de las actuales lenguas románicas. Sin embargo, se pretende ver entre esas diferencias que mantienen todos los puntos, cuáles se pueden considerar más próximos y cuáles más alejados, objetivo para el cual los dendrogramas y los gráficos MDS resultan ser óptimos.

En el Gráfico 2 de escalamiento multidimensional MDS⁶, resultado de una técnica multivariante de interdependencia, se representan en dos dimensiones las mismas relaciones vistas en el dendrograma.

cuantitativo (Marín Hernández, 2008). SPSS calcula las distancias entre todos los pares de objetos, asumiendo que cada objeto constituye un cluster. El siguiente paso es el de buscar los dos clusters más cercanos, los cuales se juntan y constituyen uno solo, el proceso se repite hasta que no quedan pares de comparación. Finalmente el resultado gráfico es un árbol de relaciones que funciona incluso si la muestra es azarosa. El dendrograma se convierte en un recurso relevante dentro de los análisis de conglomerados jerárquicos, pues facilitan tanto la interpretación de los datos como su divulgación a otros públicos, es una herramienta que brinda una capacidad explicativa por sí misma, al suministrar elementos para la conceptualización de los procesos estudiados.

⁶ El escalamiento multidimensional, más conocido como MultiDimensional Scaling (MDS), tiene sus orígenes a principios de siglo XX en el campo de la Psicología. Surge cuando se pretendía estudiar la relación que existía entre la intensidad física de ciertos estímulos con su intensidad subjetiva (Guerrero Casas y Ramírez Hurtado, 2012). El MDS es una técnica de representación espacial que trata de plasmar sobre un mapa virtual un conjunto de datos cuya posición relativa se desea analizar. Persigue como objetivo transformar los juicios de similitud o preferencia efectuados por un conjunto de individuos sobre un conjunto de estímulos susceptibles de ser representados en un espacio multidimensional. Se basa en la comparación de objetos o de estímulos, de forma que si un

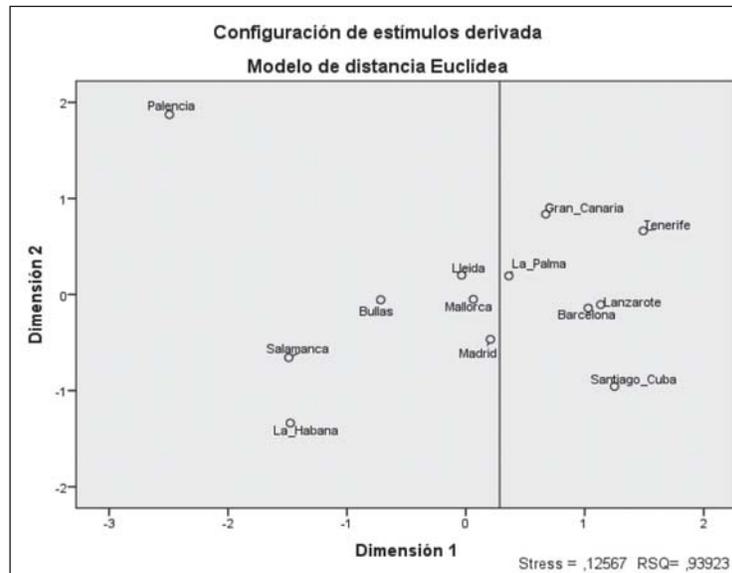


Gráfico 2. MDS de las enunciativas obtenido en SPSS⁷.

En efecto, como puede apreciarse en esta visión de los datos, en una parte del gráfico se agrupan la mayoría de variedades insulares (Tenerife, Gran Canaria, Lanzarote, La Palma y Santiago de Cuba) con otra peninsular (Barcelona), mientras que una variedad insular (La Habana) queda en el bloque básicamente peninsular. En él, Palencia queda más alejada. De todas formas, tienden a quedar –de forma bastante clara– en la parte derecha los puntos de encuesta de Canarias y Cuba, y en la parte izquierda los puntos peninsulares.

individuo juzga a los objetos A y B como los más similares entonces las técnicas de MDS colocarán a los objetos A y B en el gráfico de forma que la distancia entre ellos sea más pequeña que la distancia entre cualquier otro par de objetos. Según Guerrero Casas y Ramírez Hurtado (2012), el MDS puede ser utilizado en muchas investigaciones junto a otras técnicas multivariantes, bien como una alternativa a ellas o bien como un complemento a las mismas, a pesar de que es una técnica que sigue estando infrutilizada en muchas áreas. Entre las ventajas que señalan al MDS están que los datos en MDS pueden estar medidos en cualquier escala, mientras que en el análisis factorial deben estar medidos en escala de razón o intervalo o que en MDS pueden ser interpretadas directamente las distancias entre todos los puntos, mientras que en el análisis de correspondencias solamente pueden ser interpretadas directamente las distancias entre filas o bien entre columnas.

⁷ El Stress mínimo será el método del crecimiento más rápido o método del gradiente, método en el que es preciso el cálculo de derivadas parciales. [Un posible problema podría ser que el mínimo local no es necesariamente un mínimo absoluto, pero] solamente si la configuración obtenida tiene sentido (en términos conceptuales) se aceptará como buena, así que el hecho de que el Stress mínimo alcanzado provenga de un mínimo local o de un mínimo absoluto no afectará a la validez de los resultados en el espacio nt-configuracional. [...] No todas las medidas de proximidad deben estar

La explicación a la distribución de los datos reflejada en los gráficos 1 y 2 viene dada, sin duda, por la gran semejanza que presentan las declarativas de todas las variedades que comparten un mismo hecho fundamental, esto es, su tonema descendente, además de presentar normalmente una posposición del pico a la postónica en el pretonema.

Se demuestra así que las enunciativas no presentan grandes diferencias entre los puntos de encuesta estudiados del dominio español, igual que sucede en otras lenguas como el catalán (Carrera Sabaté, Fernández Planas y Martínez Celdrán, 2010; Fernández Planas et al., 2011; Prieto y Cabré (Eds.), 2013, para las variedades peninsulares). De hecho, los resultados perceptivos, aunque no siempre coincidan con los acústicos, señalan en la misma línea, en test de identificación de modalidad, que un final descendente (e incluso ligeramente sostenido o ascendente) se interpreta como enunciativo, sin considerar otras diferencias acústicas que puedan existir entre los puntos de encuesta incluso de lenguas diferentes, mientras que un final ascendente constituye un indicio perceptivo claro para el reconocimiento de la interrogatividad, de acuerdo con la Strong Universalist Hypothesis (Ladd, 1981), aunque también es cierto que la Nuclear Tone Hypothesis señala que la distribución de los contornos y su relación con las funciones es arbitrario y depende en cierto grado de las lenguas o de las variedades lingüísticas. Gusehoven y Chen (2000) también destacan que picos más altos, alineados con las tónicas o en el final de la frase constituyen índices universales para interpretar las emisiones como interrogativas (Fernández Planas et al., 2013; Fernández Planas et al., en prensa).

4.1.2. *Interrogativas*

Los datos correspondientes a las oraciones interrogativas estudiadas no siguen una distribución normal, igual que sucedía en las oraciones declarativas, como señala la prueba de Kolmogorov-Smirnov. La prueba de Kruskal-Wallis (Chi cuadrado=963,010; gl=13; sig=0,000) indica también que todos los puntos de encuesta presentan una configuración melódica en esta modalidad que es significa-

validadas experimentalmente (Matas, 2006: 95-100). Los resultados de la tesis doctoral de Matas (2006) demuestran que el MDS es una técnica sensible y robusta en comparación con otras técnicas análogas; que permite describir de forma válida y ajustada elementos lingüísticos para los que se han obtenido medidas acústicas; que la organización de los elementos lingüísticos (entendida en función de las posiciones de los elementos en la configuración dimensional obtenida) se realiza en función de los rasgos fonéticos relevantes; que es, comparativamente con otras técnicas, más resistente en el sentido estadístico del término; y que es enormemente versátil para adaptarse a diferentes tipos de datos. Por otra parte, el RSQ es el valor que resulta de la correlación múltiple al cuadrado entre los puntos escalares estimados y los actuales y puede interpretarse como la proporción de varianza explicada por las distancias respectivas. Se tiene un buen ajuste cuando la correlación múltiple al cuadrado es alta, como sucede en los gráficos de este tipo que incluye esta investigación.

tivamente distinta de las demás, de acuerdo con lo que se pretende como objetivo fundamental en el *Atlas Multimedia de Prosodia del Espacio Románico*. Se pretende ahora ver entre esas diferencias qué puntos presentan entre sí mayor similitud o menor distancia.

Al contrario de lo que sucede con las declarativas, en el Gráfico 3, correspondiente a las interrogativas, se observa la agrupación esperable en dos grupos de *clusters* internamente muy compactos y, a la vez, muy distantes entre sí; la única excepción es Palencia que aparece agrupada con las variedades canarias y cubanas (grupo 1) y no con las variedades peninsulares y Palma de Mallorca (grupo 2), como cabría esperar de acuerdo con las clasificaciones dialectológicas tradicionales o a partir de un criterio derivado exclusivamente de la proximidad geográfica.

La prueba U de Mann-Whitney señala que estos dos grupos mantienen una diferencia estadísticamente significativa entre sí ($Z=-28,315$; $\text{sig}=0,000$). Los grupos se diferencian por presentar un tonema ascendente uno de ellos y un tonema descendente el otro como puede comprobarse en el Atlas citado (Dorta, 2003-2015; Martínez Celadrán y Fernández Planas 2003-2015).

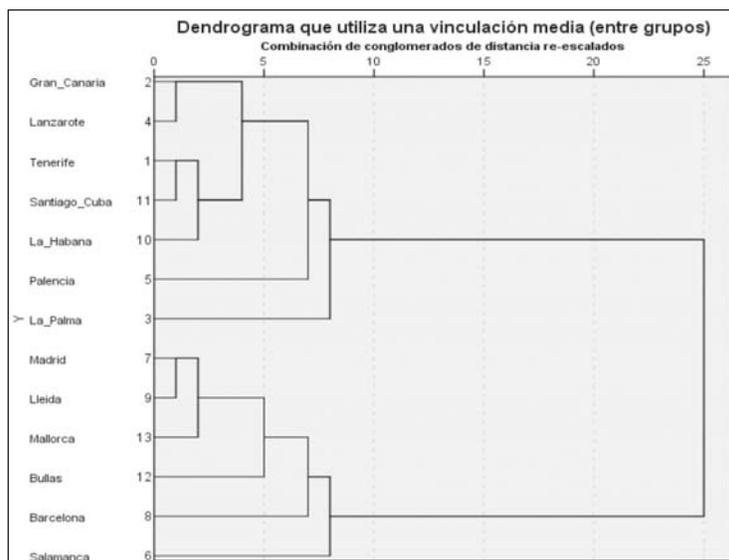


Gráfico 3. Dendrograma de las interrogativas obtenido en SPSS.

La menor distancia entre Palencia y las variedades insulares canarias y cubanas se justifica, no obstante, viendo los datos, es decir, teniendo en cuenta que dicha

variedad no presenta normalmente una configuración tonal ascendente en su final como ocurre en el resto de las variedades del segundo grupo (H%), sino descendente (L%). En efecto, tal como se puede apreciar en el Gráfico 4, Palencia tiene un final circunflejo como las variedades del grupo 1 ejemplificadas en dicho gráfico (Dorta, Ed., 2013), esto es, una canaria (Tenerife) y otra cubana (La Habana).

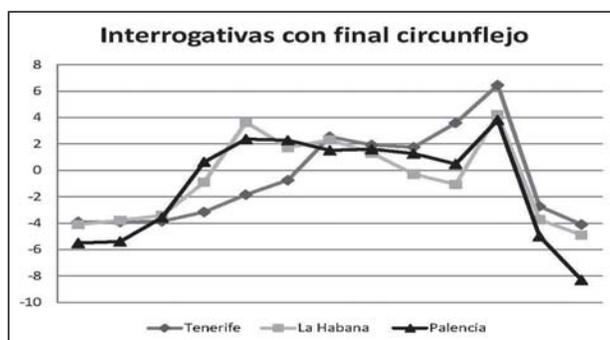


Gráfico 4. Contorno tonal en st de la oración ¿La guitarra se toca con paciencia? emitida por informantes femeninas de Tenerife (línea gris oscuro con círculos), La Habana (línea gris claro con cuadrados) y Palencia (línea negra con triángulos).

En el Gráfico 3 se puede apreciar, no obstante, que Palencia, aunque también La Palma, está más distante que el resto de las variedades canarias y cubanas, más próximas entre sí. Teniendo en cuenta la proximidad, en el primer grupo, Gran Canaria y Lanzarote se distribuyen en un mismo *cluster* debido a que hay poca distancia entre ellas, mientras que Tenerife se agrupa con Santiago y La Habana con lo cual se evidencia mayor similitud entre las variedades de ambos lados del Atlántico. En el grupo 2, las variedades menos distantes son Lleida, Madrid y Palma de Mallorca, lo que no se corresponde, obviamente, con su distribución geográfica, pero sí con los patrones entonativos de sus interrogativas absolutas.

En el MDS del Gráfico 5 se puede apreciar claramente que las variedades analizadas se distribuyen espacialmente en dos grandes bloques: el de la izquierda incluye las comprendidas en el grupo 2 del dendrograma; el de la derecha, las variedades del grupo 1.

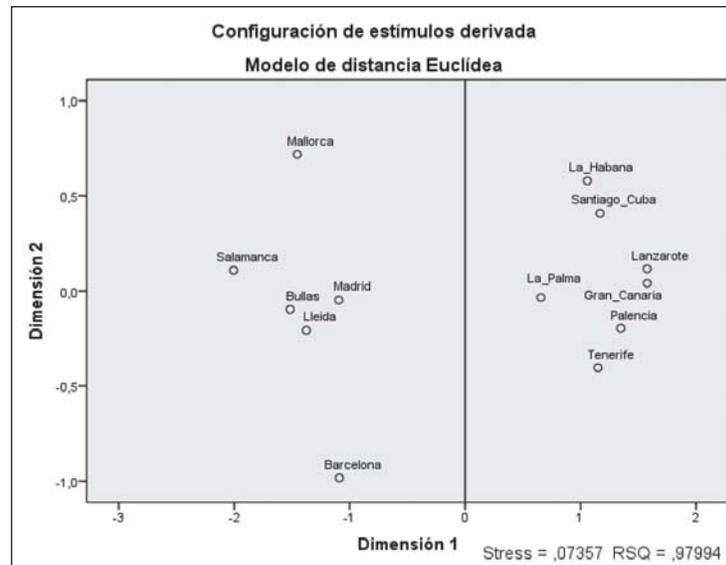


Gráfico 5. MDS de las interrogativas obtenido en SPSS.

Respecto a la modalidad, cabe destacar, como se ha hecho en trabajos anteriores tanto en el marco AMPER (Carrera Sabaté et al., 2010; Fernández Planas et al., 2011; Dorta, Ed., 2013; como en otros estudios para otras lenguas: Prieto y Cabré, 2013), que la interrogativa absoluta constituye un índice más evidente que las enunciativas para las diferencias dialectales entre los puntos de encuesta.

4.1.3. Enunciativas e interrogativas conjuntamente

La matriz de distancias a partir de la cual se pueden generar los dendrogramas y los gráficos MDS en la metodología de trabajo Calcu-Dista, teniendo en cuenta las dos modalidades juntas, es la que aparece en la Tabla III.

Tabla III. Matriz de distancias de las dos modalidades oracionales juntas.

Matriz de distancias

Caso	distancia euclídea												
	1: Tenerife	2: Gran_Canaria	3: La_Palma	4: Lanzarote	5: Palencia	6: Salamanca	7: Madrid	8: Barcelona	9: Lleida	10: La_Habana	11: Santiago_Cuba	12: Bullas	13: Mallorca
1: Tenerife	,000	,865	1,813	,687	2,587	3,789	2,734	2,049	2,699	1,681	,797	2,785	2,911
2: Gran_Canaria	,865	,000	1,775	,853	2,139	3,659	2,873	2,394	2,738	1,434	,921	2,715	2,950
3: La_Palma	1,813	1,775	,000	1,339	3,123	3,541	2,024	2,033	2,034	2,476	1,586	2,420	2,656
4: Lanzarote	,687	,853	1,339	,000	2,662	3,720	2,504	2,066	2,491	1,815	,760	2,674	2,837
5: Palencia	2,587	2,139	3,123	2,662	,000	3,827	3,784	3,449	3,530	1,916	2,503	3,398	3,562
6: Salamanca	3,789	3,659	3,541	3,720	3,827	,000	2,125	2,549	2,089	3,032	3,451	1,394	1,328
7: Madrid	2,734	2,873	2,024	2,504	3,784	2,125	,000	1,274	,888	2,783	2,410	1,273	1,094
8: Barcelona	2,049	2,394	2,033	2,066	3,449	2,549	1,274	,000	1,288	2,411	1,962	1,685	1,447
9: Lleida	2,699	2,738	2,034	2,491	3,530	2,089	,888	1,288	,000	2,832	2,493	1,317	1,140
10: La_Habana	1,681	1,434	2,476	1,815	1,916	3,032	2,783	2,411	2,832	,000	1,330	2,325	2,558
11: Santiago_Cuba	,797	,921	1,586	,760	2,503	3,451	2,410	1,962	2,493	1,330	,000	2,435	2,618
12: Bullas	2,785	2,715	2,420	2,674	3,398	1,394	1,273	1,685	1,317	2,325	2,435	,000	,807
13: Mallorca	2,911	2,950	2,656	2,837	3,562	1,328	1,094	1,447	1,140	2,558	2,618	,807	,000

Esta es una matriz de disimilitudes

El análisis dialectométrico, teniendo en cuenta las dos modalidades conjuntamente, evidencia el peso de las interrogativas en la separación de las variedades en dos grupos diferentes. Así, en el Gráfico 6 podemos comprobar que, en efecto, se vuelven a formar dos grupos de conglomerados que comprenden las mismas variedades que hemos visto en el dendrograma y en el MDS de las interrogativas (gráficos 3 y 6). Ahora bien, el hecho de que se hayan analizado las dos modalidades conjuntamente ha determinado que la distribución y las distancias entre las variedades de cada grupo hayan variado algo respecto de las que hemos visto al analizar las interrogativas por separado.

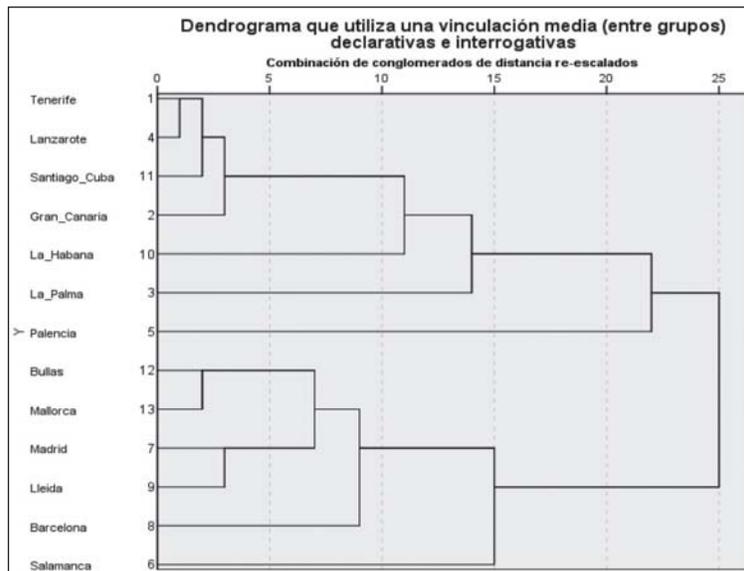


Gráfico 6. Dendrograma de las enunciativas e interrogativas obtenido en SPSS.

En efecto, en el grupo 1 de las variedades canarias y cubanas, en el que se incluye también Palencia, el punto representante del español leonés, se puede observar que esta última es la más distante –seguida de La Palma y de La Habana–, mientras que Tenerife, Lanzarote, Santiago de Cuba y Gran Canaria siguen siendo más próximas. Si se traza la línea de puntos que se ve en el Gráfico 6, el resultado es que en este grupo se forma un *cluster* de tamaño 1 con Palencia y otro de tamaño 6 formado por el resto de las variedades canarias y cubanas en el que La Palma y La Habana son más distantes respecto del resto. En el grupo 2, teniendo en cuenta la división trazada, Salamanca es el punto más distante de los demás puntos peninsulares. En este grupo, las variedades más próximas son Bullas y Mallorca y, en segundo lugar, Madrid y Lleida. Esta situación también es visible en el MDS del Gráfico 7.

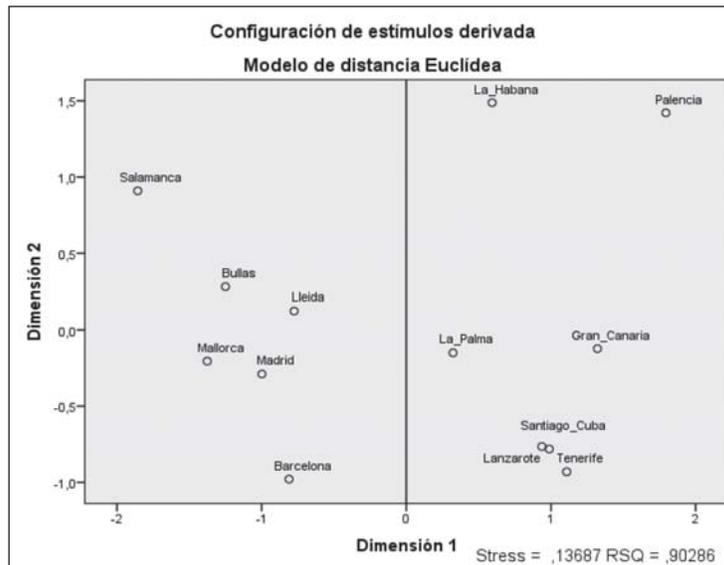


Gráfico 7. MDS de las interrogativas obtenido en SPSS.

En este gráfico se puede apreciar claramente la separación espacial de las variedades en dos grupos, de los cuales el formado por las peninsulares (salvo Palencia) y Palma de Mallorca sigue siendo más compacto. En el de la derecha, como en el dendrograma del gráfico 6, aparecen muy próximas tres islas de Canarias y Santiago de Cuba, mientras que Palencia, La Palma y La Habana son más distantes. En el lado izquierdo se aprecia claramente la gran proximidad entre Madrid, Lleida, Mallorca y Bullas, si bien Barcelona no es tan distante respecto de estas como Salamanca tal como se refleja en el dendrograma del Gráfico 6.

4.2. Análisis en DiaTech

En DiaTech se hizo un análisis de correlaciones intrasujeto (considerando a los sujetos por separado) y otro intrapunto de encuesta (para obtener la correlación de los valores de las repeticiones de cada frase en cada punto de encuesta). Una vez que se comprobó que en ambos casos la coherencia interna era alta (próxima a 1), se obtuvo el dendrograma (Gráfico 8) que muestra la formación de los dos grandes grupos ya explicados a partir de su etiquetaje prosódico superficial. Asimismo, se vuelve a comprobar que el grupo 2 es más compacto que el grupo 1, lo que implica mayor proximidad entre las variedades que lo conforman.

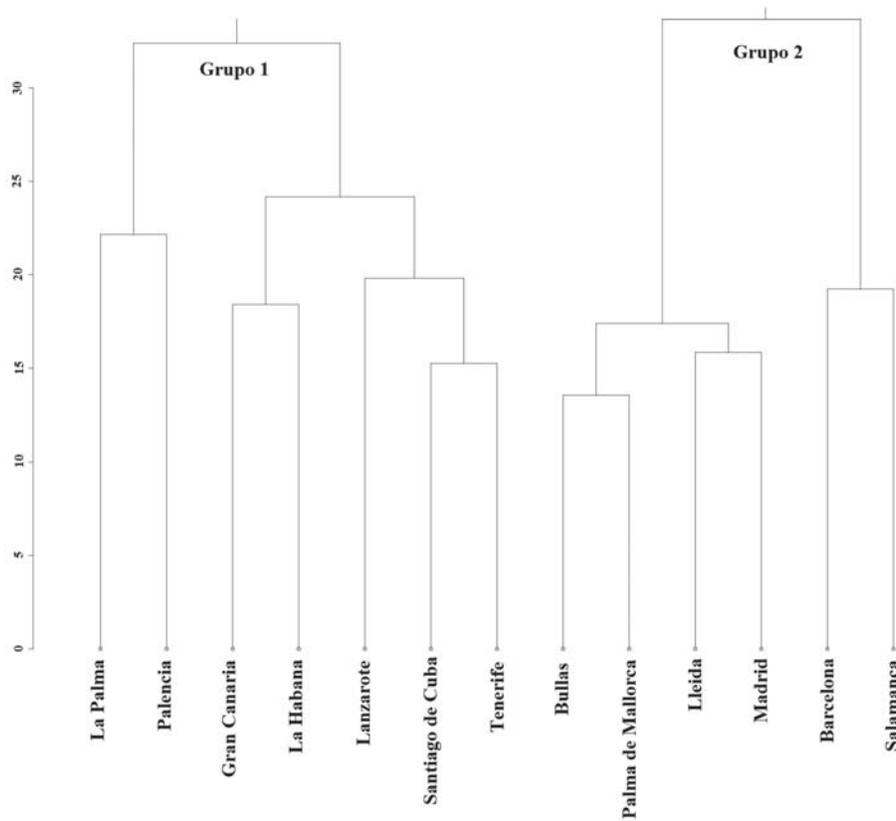


Gráfico 8. Dendrograma de declarativas e interrogativas obtenido en DiaTech.

No obstante, la distribución de las variedades ha cambiado algo respecto de las figuras anteriores. Así pues, el etiquetaje que fundamenta el análisis con esta herramienta implica cambios internos en los dos grupos de *clusters* respecto al resultado obtenido con la metodología Calcu-Dista, pero no afecta a la formación de los dos grandes grupos, ni a sus integrantes, ni a la distancia intergrupala. Se debe recordar que el análisis con DiaTech (Gráfico 8) está realizado a partir del etiquetaje superficial de los contornos y ello implica un cierto resumen de los datos empíricos, pero coincide con el anterior realizado con la metodología Calcu-Dista en la distribución de las variedades analizadas en dos grupos claramente diferenciados entre sí.

5. CONCLUSIONES

A la vista de los resultados obtenidos se pueden exponer las siguientes conclusiones respecto a los datos prosódicos estudiados tanto en la modalidad enunciativa como en la interrogativa:

1^a) Los datos en que se ha basado el presente estudio dialectométrico evidencian diferencias entre las declarativas de las diferentes variedades analizadas. No obstante, si bien tales diferencias no son del todo irrelevantes, puesto que su peso se ha dejado notar en el análisis conjunto de esta modalidad y la interrogativa, tales diferencias no parecen relevantes inequívocamente para discriminar grupos coherentes entre los puntos de encuesta estudiados.

2^a) Respecto a la hipótesis apuntada acerca de la relación histórica entre los puntos de encuesta de Canarias y América, cabe decir que, al contrario que las declarativas, las interrogativas permiten clasificar los datos estudiados en dos grandes grupos muy distantes entre sí, avalados estadísticamente: uno integrado por las variedades isleñas atlánticas de Canarias y Cuba, al que se suma Palencia por compartir una misma configuración tonal en esta modalidad, y otro constituido por el resto de las variedades peninsulares y Palma de Mallorca. Esta agrupación es consistente con lo que se conoce a propósito de la historia de la lengua española: *Dentro de las variedades que lo caracterizan [al español], las de Hispanoamérica se han relacionado históricamente con el español, de Canarias y de Andalucía distinguiéndose conjuntamente como “español meridional” o “español atlántico” en contraposición con el “español septentrional” más conservador. Las razones históricas que se aducen para explicar la relación del andaluz, canario y español americano son, por una parte, la proximidad de las fechas de la colonización de Canarias y el descubrimiento de América y, por otra, la procedencia geográfica de los colonizadores y su paso por Canarias* (Dorta, 2013:21).

3^a) Como cabía esperar, el estudio ha evidenciado que la mayor proximidad geográfica entre las variedades no motiva necesariamente una mayor proximidad

prosódica. El hecho de centrarnos exclusivamente en datos entonativos conlleva que se den casos muy llamativos como, por ejemplo, que en las interrogativas haya mayor proximidad entre Santa Cruz de la Palma y Palencia que entre la primera y el resto de las islas canarias o entre la segunda y las variedades peninsulares.

Respecto al objetivo metodológico perseguido en este trabajo cabe destacar:

1º) Los buenos resultados obtenidos respecto a la clasificación de los datos estudiados al trabajar conjuntamente las dos modalidades con los dos sistemas, además, confirma que el etiquetaje está bien hecho y que resulta ser una buena representación de los datos acústicos empíricos. El hecho de trabajar con valores numéricos, tal como hace Calcu-Dista, convierte el sistema, de alguna manera, en más sensible puesto que cualquier variación numérica puede hacer variar ligeramente el contenido interno de los bloques. Las etiquetas (tal como trabaja Dia-Tech), aunque estén hechas a partir del nivel superficial de los datos, más fonético que fonológico, ya suponen un cierto resumen de dichos datos.

2º) En relación con las metodologías de análisis, aunque las dos utilizadas presentan algunas diferencias internas en el contenido de los grandes grupos que se forman, coinciden en la distribución de los puntos de encuesta en esos grupos y en los componentes de cada uno de ellos validándose así ambas metodologías.

Agradecimientos: agradecemos la colaboración desinteresada de todos los sujetos cuyas voces se han estudiado en este trabajo y que un día las prestaron para el estudio científico de la prosodia. Este trabajo ha contado con las siguientes ayudas económicas de fondos públicos obtenidos en convocatorias competitivas: FFI2012-35998 y FFI2014-52716-P.

REFERENCIAS

- Álvarez Blanco, Rosario; Dubert, Francisco y Sousa, Xulio. (2006). Aplicación da análise dialectométrica aos datos do Atlas Lingüístico Galego. Álvarez Blanco, Rosario; Dubert, Francisco y Sousa, Xulio (Eds.): *Lingua e territorio* (461-493). Santiago de Compostela, España: Instituto da Lingua Galega - Consello da Cultura Galega.
- ArcGis Resources. (2014): Versión 10.1. <http://resources.arcgis.com/en/home/> [26/11/2014].
- Aurrekoetxea, Gotzon. (1992). Naffaroako Euskara: azterketa dialektometrikoa. *Utzaro*, 5, 59-109.
- Aurrekoetxea, Gotzon; Fernández-Aguirre, Karmele; Rubio, Jesús; Ruiz, Borja y Sánchez, Jon. (2013). DiaTech: A new tool for dialectology. *Literary and Linguistic Computing*, 28, Issue 1, 23. doi: 10.1093/lc/fqs049.
- Bauer, Roland. (2003). Sguardo dialettometrico apoyo alcune zone di transizio-

- ne dell'Italia norte-orientale (lombardo vs. Trentino vs. Veneto). Parallel X. Sguardi reciprocidad . Vicende linguistiche e cultural dell'area italoфона e germanóфона. Bombi, Raffaella y Fusco, Fabiana (Hrsg.): *Atti del Decimo Incontro italo- austriaco dei lingüista* (93-119). Udine, Italia: Forum Editrice.
- Bauer, Roland. (2005). La classificazione dialettometrica dei basiletti altoitaliani e Ladino rappresentati nell'Atlante linguistico del ladino dolomítico e dei dialetti limitófos (ALD - I). Cristina Guardiano *et al.* (Hrsg.): *Lingue, istituzioni, territorio. Riflessioni teoriche, proposte metodologiche ed esperienze di politica lingüistica* (347-365). Roma, Italia: Bulzoni.
- Beckman, Mary E. y Hirschberg, Julia. (1994). *The ToBI annotation conventions*. Manuscrito no publicado. Ohio State University y AT&T Bell Telephone Laboratories.
- Boersma, Paul y Weenink, David. (2014). Praat: doing phonetics by computer [programa]. Versión 5.4.01. Disponible en <http://www.praat.org/> [20/11/2014].
- Carrera Sabaté, Josefina; Fernández Planas, Ana María y Martínez Celdrán, Eugenio. (2010). Declaratives i interrogatives absolutes del català en el marc del projecte internacional Atlas Multimèdia de Prosòdia de l'Espai Romànic. *Caplletra*, 49, 133-167.
- Clua, Esteve. (1999). Distància lingüística i classificació de les varietats dialectal. *Caplletra*, 26, 11-36.
- Clua, Esteve. (2004). El método dialectométric: aplicación del análisis multivariante a la clasificación de las variedades del catalán. M.P. Perea (Ed.): *Dialectología y recursos informáticos* (59-88). Barcelona, España: Universitat de Barcelona.
- Contini, Michel. (1992). Vers une géoprosodie. *Actes du «Nazioarteko Dialektologia Biltzarra Agiriak»* (83-109). Bilbao, España: Real Academia de la Lengua Vasca.
- Contini, Michel; Lai, Jean Pierre y Romano, Antonio. (2002). La géolinguistique à Grenoble: de l'AliR à AMPER. *Revue Belge de Philologie et d'Histoire*, fasc. 3, 80, 931-941.
- Contini, Michel; Lai, Jean Pierre; Romano, Antonio y Rouillet, Stefania. (1998). Vers un atlas prosodique parlant des variétés romanes. Bouvier, Jean-Claude *et al.* (Eds.): *Mélanges offerts à X. Ravier* (73-84). Toulouse, Francia: Université de Toulouse–Le Mirail.
- Contini, Michel; Lai, Jean Pierre; Romano Antonio; Rouillet Stefania, Moutinho, Lurdes de Castro; Coimbra, Rosa Lidia; Pereira Bendiha, Urbana y Secca Ruivo, Suzana. (2002). Un projet d'atlas multimèdia prosodique de l'espace roman. *Proceedings of the International Conference Speech Prosody 2002* (227-230). Aix-en-Provence, Francia: Laboratoire Parole et Langage.
- D'Andrés Díaz, Ramón; Álvarez-Balbuena García, Fernando y Suárez Fernández,

- Xosé Miguel. (2007). Proxecto ETLEN para o estudio dialectográfico e dialectométrico da zona Eo-Navia, Asturias: fundamentos teóricos. *Actas VII Congreso Internacional de Estudos Galegos: mulleres en Galicia: Galicia e os outros pobos da península* (749-759). A Coruña: Edicións do Castro.
- Dorta Luis, Josefa. (2009). La investigación geolingüística en Canarias: hacia un atlas de prosodia. Corbella, Dolores y Dorta, Josefa (Eds.): *La Investigación dialectológica en la actualidad* (347-371). Santa Cruz de Tenerife, España: Agencia Canaria de Investigación.
- Dorta, Josefa (Ed.). (2003-2015). *AMPER-CAN. Atlas Multimedia de la Prosodia del Espacio Románico*. Disponible en <http://ampercan.webs.ull.es/>.
- Dorta, Josefa (Ed.). (2013): *Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela*. Santa Cruz de Tenerife, España: La Página Ediciones.
- Elvira-García, Wendy; Roseano, Paolo; Fernández Planas, Ana María y Martínez Celdrán, Eugenio. (2014). Eti-ToBI: una herramienta para la transcripción prosódica automática con etiquetas Sp_ToBI en Praat. Comunicación presentada en el *VI Congreso Internacional de Fonética Experimental* del 5-7 de noviembre de 2014. Universitat de València, Valencia, España.
- Estebas Vilaplana, Eva y Prieto, Pilar. (2009). La notación prosódica en español. Una revisión del Sp_ToBI. *Estudios de Fonética Experimental*, 18, 263-283.
- Fernández Planas, Ana María. (2005). Datos generales del proyecto AMPER en España. *Estudios de Fonética Experimental*, 14, 13-27.
- Fernández Planas, Ana María y Martínez Celdrán, Eugenio. (2003). El tono fundamental y la duración: Dos aspectos de la taxonomía prosódica en dos modalidades de habla (enunciativa e interrogativa) del español. *Estudios de Fonética Experimental*, 12, 165-200.
- Fernández Planas, Ana María; Roseano, Paolo; Martínez Celdrán, Eugenio y Romera Barrios, Lourdes. (2011). Aproximación al análisis dialectométrico de la entonación en algunos puntos del dominio lingüístico catalán. *Estudios de Fonética Experimental*, 20, 141-178.
- Fernández Planas, Ana María; Roseano, Paolo; Elvira-García, Wendy; Carrera Sabaté, Josefina; Román Montes de Oca, Domingo y Martínez Celdrán, Eugenio. (2013). Impressions perceptives sobre dades del català, de l'italià i del sard en el marc de l'Atles Multimèdia de Prosòdia e l'Espai Romànic. Comunicación presentada en el *V Workshop sobre Entonación del Catalán i Cat_Tobi* en junio de 2013. Universitat Pompeu Fabra, Barcelona, España.
- Fernández Planas, Ana María; Roseano, Paolo; Elvira-García, Wendy; Carrera Sabaté, Josefina; Román Montes de Oca, Domingo y Martínez Celdrán, Eugenio (en prensa). From a perceptual point of view, is there prosodic continuity between languages in contact? Congosto, Yolanda y Morgenthaler, Laura (Coords). Amsterdam, Holanda: Issues in Hispanic and Lusophon Linguistics. John Benjamins.

- Goebel, Hans. (1981). Eléments d'analyse dialectométrique (avec application à l' AIS). *Revue de Linguistique Romane*, 45, 349-420.
- Goebel, Hans. (1987). Encore un golpe de oeil dialectométrique sur las Tableaux phonétiques diciembre patois suizos permanecer (TPPSR). Deux analyses interponctuelles: parquet polygonal te Treilles triangulaire. *Vox románica*, 46, 91-125.
- Goebel, Han. (1993). Die dialektale Gliederung Ladinien aus der Sicht der ladinno. Eine Pilotstudie zum Problem der geolinguistischen «Mental Maps». *Ladina*, 17, 59-95.
- Goebel, Hans. (2003). Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur. *ER*, 25, 59-120.
- Goebel, Hans. (2013). La dialectometrización del ALPI: rápida presentación de los resultados. Casanova, Emili y Calvo, Cesáreo (Eds.): *Actas del XXVI Congreso Internacional de Lingüística y Filología Románicas*, 6, (143-154). Berlín, Alemania: Gruyter.
- Goebel, Hans y Schiltz, Guillaume. (1997). Dialectometrical compilation of CLAE 1 and CLAE 2. Isoglosses and dialect integration. Viereck, Wolfgang; Ramisch, Heinrich; Händler, Harald y Marx, Christian (Edd.): *The computer developed linguistic Atlas of England*, 2, (13-21). Tübingen, Alemania: Niemeyer.
- Guerrero Casas, Flor María y Ramírez Hurtado, José Manuel. (2012). El análisis de escalamiento multidimensional: una alternativa y un complemento a otras técnicas multivariantes. *La Sociología en sus escenarios*, Revista electrónica de la Universidad de Antioquía (Colombia), 25. Disponible en <http://aprendeenlinea.udea.edu.co/revistas/index.php/ceo/article/view/11450>.
- Gussenhoven, Carlos y Chen, Aoju. (2000). Universal and language-specific effects in the perception of question intonation. Yuan, Baozong; Huang, Taiyi y Tang, Xiaofang (Eds.): *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP)* (91-94). Beijing, China: China Military Friendship Publish.
- Heeringa, Wilbert y Nerbonne, John. (2001). Dialect Areas and Dialect Continua. Sankoff, David; Labov, William y Kroch, Anthony (Eds.): *Language Variation and Change*, 13, (375-400). Cambridge, Inglaterra: Cambridge University Press.
- Hermes, Dik J. (1998a). Auditory and visual similarity of pitch contours. *Journal of Speech, Language and Hearing Research*, 41, 63-72.
- Hermes, Dik J. (1998b). Measuring the Perceptual Similarity of Pitch Contours. *Journal of Speech, Language and Hearing Research*, 41, 73-82.
- Ladd, D. Robert. (1981). On Intonational Universals. Myers, T. et al. (Eds.): *The*

- Cognitive Representation of Speech* (389-397). Amsterdam, Holanda: North Holland Publishing.
- Ladd, D. Robert. (2008). *Intonational phonology*. Cambridge, Inglaterra: Cambridge University Press.
- López Bobo, María Jesús; Muñiz Cachón, Carmen; Díaz Gómez, Liliana; Corral blanco, Norberto; Brezmes Alonso, David y Alvarellos Pedrero, Mercedes. (2007). Análisis y representación de la entonación. Replanteamiento metodológico en el marco del proyecto AMPER. Dorta, Josefa (Ed.): *La prosodia en el ámbito lingüístico románico* (17-34). Santa Cruz de Tenerife, España: La Página Ediciones.
- Marín Hernández, Juan José. (2008). Los análisis clusters de tipo jerárquico y los dendrogramas. Una visión para la triangulación metodológica en los estudios comparativos regionales. Cairo, Heriberto y De Sierra, Gerónimo (compiladores): *América Latina, Una y diversa: Teorías y métodos para su análisis*. San José, Costa Rica: Editorial Alma Máter, Universidad de Costa Rica, Universidad Complutense y Universidad de la República.
- Martínez Celdrán, Eugenio y Fernández Planas, Ana María (Eds.) (2003-2015). *AMPER-CAT. Atlas Multimedia de la Prosodia del Espacio Románico*. Disponible en http://stel.ub.edu/labfon/amper/cast/index_ampercat.html.
- Martínez Celdrán, Eugenio y Fernández Planas, Ana María. (2003). Taxonomía de las estructuras entonativas de las modalidades declarativa e interrogativa del español estándar peninsular estándar según el modelo AM en habla de laboratorio. Herrera, Esther y Martín, Pedro (Eds.): *La tonía: dimensiones fonéticas y fonológicas* (267-294). México D.F. México: El Colegio de México.
- Matas Crespo, Josep. (2006). *La técnica del Escalamiento Multidimensional en el vocalismo: un análisis comparativo*. Tesis Doctoral. Barcelona, España: Universitat de Barcelona.
- Moutinho, Lurdes de Castro; Coimbra, Rosa Lidia; Rilliard, Albert y Romano, Antonio (2011). Measure de la variation prosodique diatopique en portugais européen. *Estudios de Fonética Experimental*, 20, 33-55.
- Pamies Bertrán, Antonio; Fernández Planas, Ana María; Martínez Celdrán, Eugenio; Ortega Escandell, Alicia y Amorós Céspedes, Mari Cruz. (2001). Umbrales tonales en español peninsular. *Actas del II Congreso de Fonética Experimental* (272-278). Sevilla, España: Universidad de Sevilla.
- Pierrehumbert, Janet. (1980). *The phonetics and phonology of English intonation*. Doctoral dissertation, Boston, Estados Unidos: Massachusetts Institute of Technology.
- Polanco, Lluís. (1992). Lengua y dialecto: una aplicación dialectométrica a la lengua catalana. *Miscelánea Sanchis Guarner*, 3, 5-28.

- Prieto, Pilar y Roseano, Paolo (Eds.). (2010). *Transcription of Intonation of the Spanish Language*. München, Alemania: Lincom Europa.
- Prieto, Pilar y Cabré, Teresa (Eds.). (2013). *L'entonació dels dialectes catalans*. Barcelona, España: Publicacions l'Abadia de Montserrat.
- Rietveld, Toni y Gussenhoven, Carlos. (1985). On the relation between pitch excursion size and prominence. *Journal of Phonetics*, 13, 299-308.
- Rilliard, Albert y Lai, Jean Pierre. (2008). Outils pour le calcul et la comparaison prosodique dans le cadre du projet AMPER, l'exemple des variétés Occitane et Sarde. Turculeț, Adrian (Org.): *La variation diatopique de l'intonation dans le domaine roumain et roman* (217-229). Iași, Rumanía: Editura Universităţii Al. I. Cuza.
- Romano, Antonio. (1999, vol. II). Analyse des structures prosodiques des dialectes et de l'italien régional parlés dans le Salento (Italie): approche linguistique et instrumentale. Tesis doctoral. Grenoble, Francia: Université Stendhal.
- Romano, Antonio. (2003, vol. I). Un projet d'Atlas multimédia prosodique de l'espace roman (AMPER). Sánchez Miret, Fernando (Ed.): *Acti del XXIII CIL-FR* (279-294). Tübingen, Alemania: Niemeyer.
- Romano, Antonio y Contini, Michel. (2001). Un progetto di Atlante geopro-sodico multimediale delle varietà linguistiche romanze. Magno Caldognetto, Emanuela e Cosi, Piero (a cura di): *Multimodalità e Multimedialità nella Comunicazione, Atti delle XI Giornate di Studio del "Gruppo di Fonetica Sperimentale" dell'Ass. Italiana di Acustica* (121-126). Padova, Italia: Unipress.
- Romano, Antonio e Interlandi, Grazia M. (2005). Variabilità geo-socio-prosodica: dati linguistici e statistici. *Géolinguistique*, Hors serie n.3, 259-280.
- Romano, Antonio y Miotti, Renzo. (2008). Distancias prosódicas entre variedades románicas. Turculeț, Adrian (Org.): *La variation diatopique de l'intonation dans le domaine roumain et roman* (231-249). Iași, Rumanía: Editura Universităţii Al. I. Cuza.
- Romano, Antonio; Lai, Jean Pierre y Rouillet, Stefania. (2005). La méthodologie AMPER. *Géolinguistique*, Hors série 3, 1-5.
- Romano, Antonio; Contini, Michel; Lai, Jean Pierre y Rilliard, Albert. (2011). Distancias prosódicas entre variedades románicas en el marco del proyecto AMPER. *Revista Internacional de Lingüística Iberoamericana*, 1, 17, 13-26.
- Roseano, Paolo. (2012). *La prosòdia del friulà en el marc de l'Atlas Multimèdia de Prosòdia de l'Espai Romànic*. Tesis doctoral. Barcelona, España: Universitat de Barcelona.
- Roseano, Paolo y Fernández Planas, Ana María. (2013). Transcripció fonètica i fonològica de l'entonació: una proposta d'etiquetatge automàtic. *Estudios de Fonètica Experimental*, 22, 275-332.
- Roseano, Paolo; Fernández Planas, Ana María y Martínez Celdrán, Eugenio. (2014). El etiquetaje entonativo autosegmental métrico en el marco del Atlas

- Multimedia de Prosodia del Espacio Románico. Congosto, Yolanda; Salvador, Antonio y Montero Curiel, María Luisa (Eds.): *Actas del V Congreso Internacional de Fonética Experimental*. Cáceres, España: Universidad de Extremadura.
- Roseano, Paolo; Elvira-García, Wendy; Fernández Planas, Ana María y Martínez Celdrán, Eugenio. (En preparación). El sistema Calcu-Dista, una herramienta para trabajar en dialectometría en el seno del *Atlas Multimedia de Prosodia del Espacio Románico*.
- Saramago, João (2002): Diferenciação lexical interpontual nos territórios galego e português (Estudo dialectométrico aplicado a materiais galegos do ALGa). Álvarez, Rosario; Dubert García, Francisco e Sousa Fernández, Xulio (eds.): *Dialectoloxía e léxico*. Santiago de Compostela, Instituto da Lingua Galega / Consello da Cultura Galega, Sección de Lingua, 41-68.
- Séguy, Jean. (1971). La relation entre la distance spatiale et la distance lexical. *Revue de Linguistique Romane*, 35, 335-357.
- Séguy, Jean. (1973). La dialectométrie dans l'Atlas lingüística de la Gascuña. *Revue de Linguistique Romane*. 37, 1-24.
- Sousa, Xulio. (2006). Análise dialectométrica das variedades xeolingüísticas galegas. Y Encontro de estudos dialectológicos. Rolão Bernardo, Maria Clara y Mateus Montenegro, Helena (Eds.): *Actas do I Encontro de Estudos Dialectológicos* (345-362). Ponta Delgada, Portugal: Instituto Cultural de Ponta Delgada.
- Verlinde, Serge. (1988). La dialectométrie et la détection des zones dialectales: l'architecture dialectale de l'Est de la Belgique romane. *Revue de Linguistique Romane*, 51, 151-172.