

DOS DÉCADAS DE DIALECTOMETRÍA ENTONATIVA

Paolo Roseano
Universitat de Barcelona
paolo.roseano@ub.edu

Abstract

This article summarizes the most important developments of intonational dialectometry during the last two decades (1998-2016). After a smooth start at the very end of the 20th century, intonational dialectometry has become more and more popular during the last ten years. During the last decade researchers interested in Romance languages have developed a few different computer programs and packages that can be used to carry out dialectometrical analysis of intonational data. The most recent application of this kind, called ProDis, was developed at the Phonetics Laboratory of the University of Barcelona.

Keywords: dialectometry, intonation, prosody, ProDis

Introducción

La preocupación por la necesidad de medir de manera objetiva las distancias entre variedades lingüísticas surgió en los años 70 del siglo XX y ha llevado al nacimiento de la disciplina conocida como dialectometría (v. entre otros, Séguy, 1971; Goebel, 1982), que consiste en aplicar varios tipos de técnicas de análisis estadístico a bases de datos muy extensas (normalmente atlas lingüísticos). Las bases de datos sobre las cuales se suelen aplicar las técnicas dialectométricas incluyen datos fonético-fonológicos segmentales, morfológicos, léxicos o sintácticos, pero no incluyen información acerca de la entonación. De hecho, sólo en la última década se han creado las bases de datos entonativas – los atlas prosódicos como AMPER (Contini, 1992; Contini, Rouillet, Romano y Lai, 2003) o IARI (Prieto, Roseano y Borràs-Comes, 2010-2014) – que constituyen el material imprescindible para el análisis geolingüístico de la entonación. Este retraso en la creación de bases de datos prosódicos ha comportado que hoy en día sean aún muy escasos los estudios que aplican métodos estadísticos para valorar las diferencias y semejanzas entre la entonación de diferentes dialectos o lenguas (Hermes 1998a,b; Romano, 1999; Moutinho, Coimbra, Rilliard y Romano, 2011; Rilliard y Lai, 2008; Romano y Miotti, 2008; Fernández Planas, Roseano, Martínez Celdrán y Romera, 2011; Sullivan, 2011; Romano, Rilliard, Lai y Contini, 2011; Roseano, 2012; Fernández Rei, Moutinho y Coimbra, 2013; Prieto y Cabré, 2013). Este artículo, presenta una breve historia de la evolución de la que podríamos definir la dialectometría entonativa. El objetivo es, por lo tanto, presentar el estado de la cuestión del desarrollo de técnicas e instrumentos para el análisis dialectométrico de datos prosódicos.

1. La dialectometría tradicional, sus datos y técnicas de análisis

Los atlas lingüísticos representan una fuente importante de informaciones sobre los rasgos fonéticos, morfológicos, sintácticos y, sobre todo, léxicos de una lengua y de sus variedades. Además, constituyen un punto de partida para la detección de similitudes entre dialectos, lo que permite clasificarlos y agruparlos. Este proceso de clasificación se ha desarrollado, en los estudios tradicionales de dialectología, de forma cualitativa, es decir que cada investigador selecciona una serie de rasgos relevantes, les

otorga menor o mayor importancia según criterios que el investigador mismo (o la tradición) considera importantes, y los utiliza para establecer las agrupaciones dialectales.

En los años 70 del siglo XX surgió una nueva tendencia metodológica en el ámbito de los estudios dialectológicos, que se llamó dialectología cuantitativa o también dialectometría, término propuesto por Séguy (1971). Esta disciplina, según la definición de Goebel (1981: 349), representa una alianza entre la geolingüística y la taxonomía numérica como rama de la matemática. La dialectometría ofrece algunas ventajas claras en comparación con la dialectología tradicional. En primer lugar, permite gestionar una cantidad de datos muy grande con un esfuerzo relativamente limitado. Justamente el hecho de trabajar con bases de datos extensas hace posible que las conclusiones a las que llega el análisis dialectométrico tengan un valor estadístico que va más allá de la intuición de un investigador. En segundo lugar, las técnicas de análisis dialectométrico están menos expuestas al riesgo de apriorismos, ya que no otorgan una importancia menor o mayor a las diferentes variables que conforman la base de datos, sino que cada una de ellas tiene el mismo peso que las otras. Aun así, hay que recordar que, justamente por este motivo, estos métodos no han estado exentos de críticas desde la dialectología tradicional, ya que no tienen en cuenta que ciertas diferencias lingüísticas son cualitativamente más relevantes que otros o también porque puede ser discutible la medida de similitud a partir de la cual se establecen las distancias, tal como apunta Clua (1999). Una última ventaja técnica de las técnicas dialectométricas en comparación con la dialectología tradicional consiste en la forma de presentación de las correlaciones estadísticas entre los datos. De hecho, para evitar que la interpretación de los resultados del tratamiento estadístico de los datos lingüísticos (mediante análisis como el *cluster analysis* y el *multidimensional scaling*) se convirtiera en una barrera infranqueable para muchos lingüistas, a lo largo de las últimas cuatro décadas se han ensayado diversas formas de representación gráfica, entre ellas los dendrogramas y los mapas MDS, que actualmente se utilizan de manera habitual en este tipo de trabajos.

Los análisis dialectométricos tradicionales se realizan exclusivamente a partir de bases de datos de tipo léxico, fonético segmental y morfológico (entre los muchos ejemplos, podemos citar para el catalán Clua, 2004; Perea, 2010; Valls, Nerbonne, Prokic, Wieling, Clua y Lloret, 2012; Goebel, 2013, entre otros). Solo a partir de los primeros años del siglo XXI se empezó a disponer de las bases de datos suprasegmentales que, a partir de la primera aproximación de Romano (1999), se analizaron mediante técnicas dialectométricas. Desde el punto de vista metodológico, como se verá más adelante, hay que destacar que las bases de datos suprasegmentales difieren radicalmente de las bases de datos dialectales tradicionales, ya que las segundas son alfabéticas (p.e. las transcripciones de la pronunciación de una palabra), mientras que las primeras son numéricas (pe los valores numéricos, en Hz o st, del F0 en diferentes puntos de un contorno entonativo). Esta diferencia en el tipo de datos, aparentemente banal, es en realidad crucial para entender por qué razón en el marco de la dialectología entonativa con datos numéricos no se ha podido utilizar ninguna de las herramientas dialectométricas más conocidas, es decir *Visual DialectoMetry* (Haimlerl, 2006), *Gabmap* (Nerbonne, Colen, Gooskens, Kleiweg y Leinonen, 2011) y *DiaTech* (Aurrekoetxea Fernandez-Aguirre, Rubio, Ruiz y Sánchez, 2013).

En definitiva, el proceso de la dialectometría tradicional se puede representar esquemáticamente en las tres fases que aparecen en la Figura 1: la de recogida de datos

alfabéticos, la de creación de una matriz de distancias numéricas a partir de los datos alfabéticos, y la de transformación de la matriz en cuestión en una (o más de una) representación gráfica de más fácil interpretación.

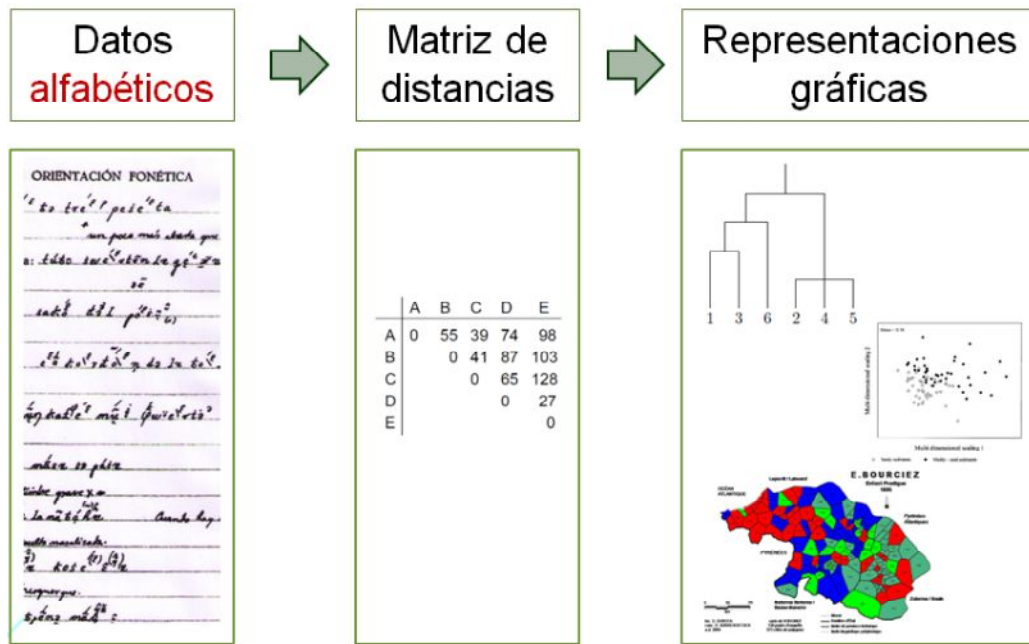


Figura 1. Fases de análisis en dialectometría tradicional

2. La dialectometría prosódica: objetivo general

Los estudios de prosodia (y su variación dialectal) se pusieron de moda solo en las últimas décadas, y las grandes bases de datos de ese tipo se han ido recogiendo solo en los últimos años. Cuando los entonólogos quisieron aplicar técnicas dialectométricas a sus bases de datos (que contenían sobre todo datos numéricos), se encontraron con que no existían instrumentos adecuados, ya que –tal y como se ha adelantado en la Sección anterior– los programas de análisis dialectométrico que existían estaban pensados para analizar solo (o casi solo) datos alfabéticos. Por esta razón, uno de los objetivos metodológicos más urgentes para los investigadores que se ocupaban de dialectología entonativa era el de crear un instrumento dialectométrico adecuado, que fuera capaz de reproducir las tres etapas de la dialectología tradicional, pero teniendo en cuenta que los datos de base son numéricos. En otras palabras, el proceso que se quería implementar en el marco de la dialectometría entonativa se puede representar esquemáticamente en las tres fases que aparecen en la Figura 2 y que son muy parecidas a las que se han contemplado en la Figura 1: la de recogida de datos numéricos (básicamente, de valores de F0 en distintos puntos de una curva entonativa), la de creación de una matriz de distancias numéricas a partir de esos datos, y la de transformación de la matriz en una representación gráfica transparente.

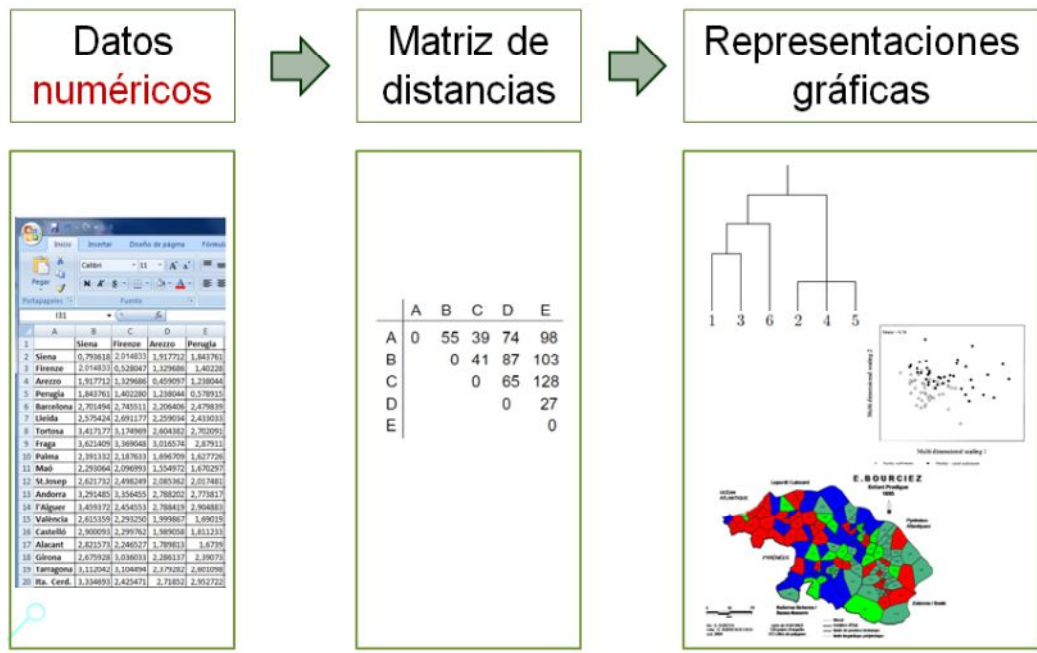


Figura 2. Fases de análisis en dialectometría entonativa

3. La dialectometría entonativa: un esquema de su historia

La dialectometría entonativa es, de hecho, una disciplina con una historia relativamente breve, que empieza a finales de los años noventa y que se puede representar con un esquema como el de la Figura 3, que se irá comentando y desarrollando a lo largo de los próximos apartados.

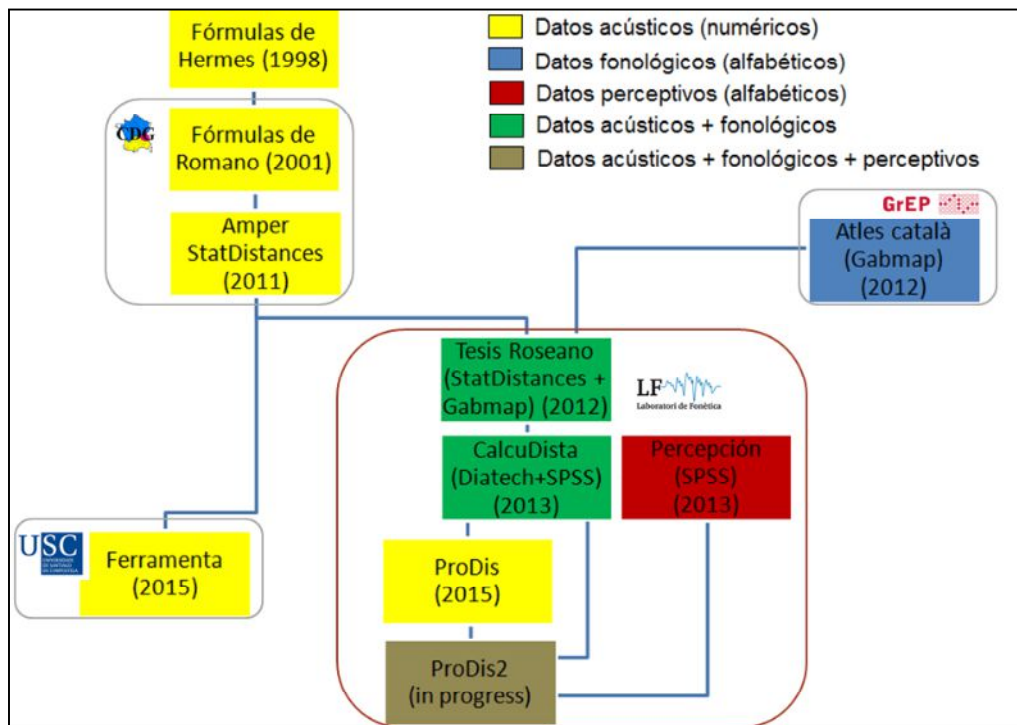


Figura 3. Fases históricas de la dialectometría entonativa. Los colores indican los diferentes tipos de datos que se someten a análisis dialectométrico y los recuadros identifican las aportaciones de diferentes grupos de estudio.

4. La dialectometría entonativa: sus bases de datos

Para entender el camino que ha seguido el desarrollo de los programas dialectométricos entonativos es necesario tener en cuenta las bases de datos entonativas de las que se dispone hoy en día (en concreto para las lenguas románicas, que son las únicas a las que se han aplicado las técnicas de dialectometría entonativa). Las bases de datos en cuestión son, fundamentalmente, dos, o, si queremos, de dos familias: la familia AMPER (Contini, 1992) y la familia IARI (Prieto et ál., 2010-2014; Frota y Prieto, 2015). Ambos proyectos no solo son similares en sus objetivos, sino también en su organización y en sus resultados. Sin embargo, muestran diferencias en los datos que se analizan. En lo que se refiere a la organización, ambos proyectos tienen un comité central de dirección y grupos de investigación para cada idioma (y, a veces, en el caso de lenguas geográficamente extensas, grupos de investigación regionales). Los datos se ponen a disposición del público en línea en el sitio web central de cada proyecto, con muestras de todos los idiomas, y en las páginas web de los grupos de investigación que recopilan datos de lenguajes específicos (como Martínez Celdrán y Fernández Planas, 2003-2016a y Prieto y Cabré, 2007-2012 catalán; Prieto y Sichel-Bazin 2014 por el occitano; Prieto y Roseano, 2009-2013 y Martínez Celdrán y Fernández Planas, 2003-2016b de español; Frota y Cruz, 2012-2015 y de Castro Moutinho, 2008-2013 para el portugués; Fernández Rei, 2016 para el gallego; Roseano y Fernández Planas, 2009-2013 para friulano, Turculeţ, 2010-2013 para el rumano, entre otros).

La diferencia quizás más importante entre AMPER y IARI es el método de análisis que, a su vez, determina el tipo de información que está contenida en las bases de datos mencionadas anteriormente. Mientras que en la familia IARI el método predominante de análisis es auditivo y fonológico, en la familia AMPER el enfoque es principalmente fonético y numérico. Eso tiene consecuencias importantes sobre el tipo de datos que constituyen la base de datos final, la que se puede utilizar para análisis de tipo dialectométrico. Mientras que las bases de datos en línea IARI suelen ofrecer el contorno F0 y el etiquetaje fonológico de un solo ejemplo de un gran número de modalidades oracionales por punto de encuesta (desde 18 a 47, dependiendo del idioma), las bases de datos AMPER ofrecen los valores numéricos (en Hz st, ms y dB) de la prosodia de varias grabaciones (de 378 a 567, dependiendo de la lengua) de dos tipos de oraciones (declarativas neutras e interrogativas totales neutras). La Figura 4 contiene un ejemplo de datos entonativos que se ofrecen en los atlas en línea en español de la familia IARI (Prieto y Roseano, 2009-2013).

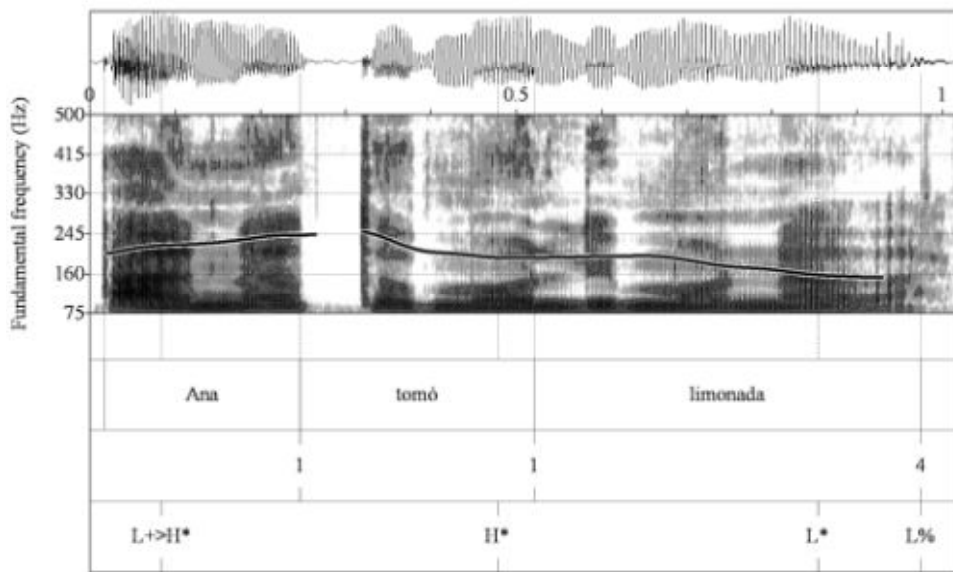


Figura 4. Ejemplo de datos en la familia IARI

Por otra parte, las bases de datos AMPER contienen valores numéricos seleccionados que resumen la prosodia de cada oración; básicamente, para cada sílaba de la frase, la base de datos contiene tres valores F0 (en Hz y st), tres valores de intensidad (en dB) y la duración (en ms). Basándose en estos valores numéricos, un transcriptor recientemente desarrollado *ad hoc* llamado *AmperEti* (Roseano y Fernández Planas, 2013) puede añadir una transcripción entonativa fonética estrecha del contorno de cada frase. Esto significa que, desde hace muy pocos años, la base de datos AMPER puede contener, además de sus valores fonéticos prosódicos numéricos habituales, también un conjunto de transcripciones fonéticas (es decir, de datos alfabéticos) de los contornos entonativos. La Figura 5 resume el proceso de análisis en el marco AMPER, en el que se puede destacar la presencia de una base de datos numérica.

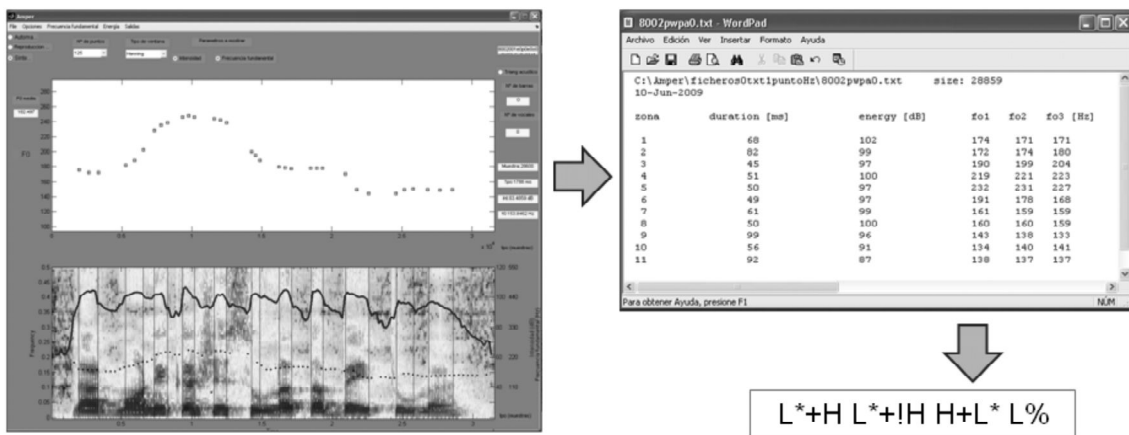


Figura 5. Ejemplo de datos en la familia AMPER.

En definitiva, mientras que la base de datos IARI contiene datos alfabéticos (es decir, las etiquetas entonativas) que se pueden intentar analizar con los instrumentos de la dialectometría tradicional, la base de datos AMPER contiene típicamente datos

numéricos para cuyo análisis los instrumentos como *VisualDialectoMetry*, *Gabmap* o *DiaTech* no resultan adecuados.

5. La dialectometría entonativa numérica: su origen

Los primeros intentos de calcular la distancia prosódica entre dos curvas entonativas se deben a Hermes (1998a,b), que propuso las primeras fórmulas que tenían esa finalidad. Su propuesta, sin embargo, no llegó a implementarse inmediatamente en el sector dialectológico. Se tuvo que esperar algunos años para que eso sucediera. A principios del siglo XXI, de hecho, es cuando se empiezan a aplicar las fórmulas de Hermes a bases de datos de variación diatópica de la entonación. En esta primera fase el centro impulsor de la dialectometría entonativa es el grupo de trabajo que tiene como punto de referencia en Centro de Dialectología de Grenoble. En concreto, es con los trabajos de investigación relacionados con la tesis doctoral de Romano (1999) que se empieza a vislumbrar la posibilidad alentadora de aplicar técnicas dialectométricas a la base de datos prosódicos que se estaba empezando a construir en aquellos años: AMPER (Contini, 1992; Contini et ál., 2003).

Los dos programas que se utilizan en esta primera fase utilizan, básicamente, las fórmulas de Hermes (1998a,b) o adaptaciones de las mismas. El primero de ellos, conocido como *Amper StatDistances* fue creado por Albert Rillard y Jean-Pierre Lai en el marco del proyecto AMPER (Rillard y Lai, 2008; Moutinho et ál., 2011), específicamente con el objetivo de dialectometrizar los datos numéricos de la F0 de las frases que se recogen de acuerdo con la metodología del proyecto. Este aplicativo, a partir de los datos prosódicos que corresponden a tres valores de F0 (en Hz) en cada vocal de la frase (inicial, central y final), más un valor de intensidad en cada una de ellas, ofrece una representación gráfica de las proximidades entre dialectos y de las agrupaciones de los mismos a partir de un dendrograma y un mapa MDS. Las figuras 6 y 7 contienen un ejemplo de dendrograma y mapa MDS, ambos realizados con *Amper StatDistances* y adaptados de Roseano (2012).

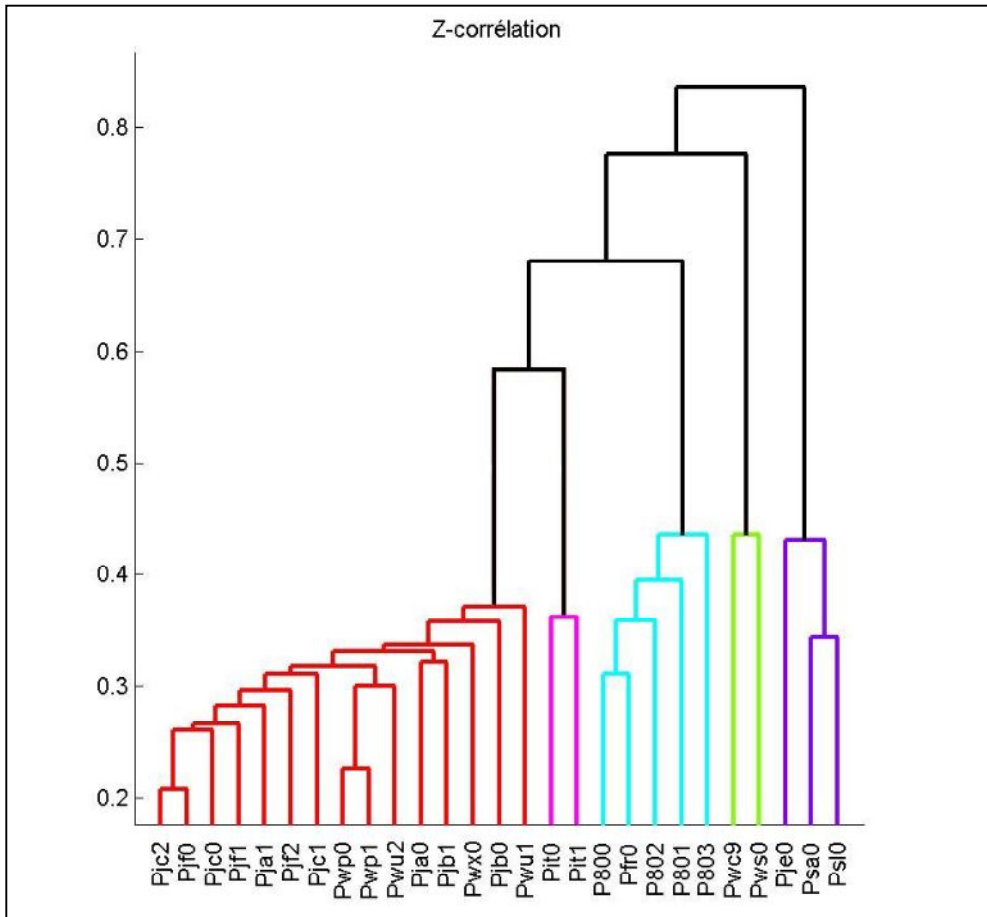


Figura 6. Dendrograma que representa las distancias entonativas entre distintas variedades románicas realizado con *Amper StatDistances*

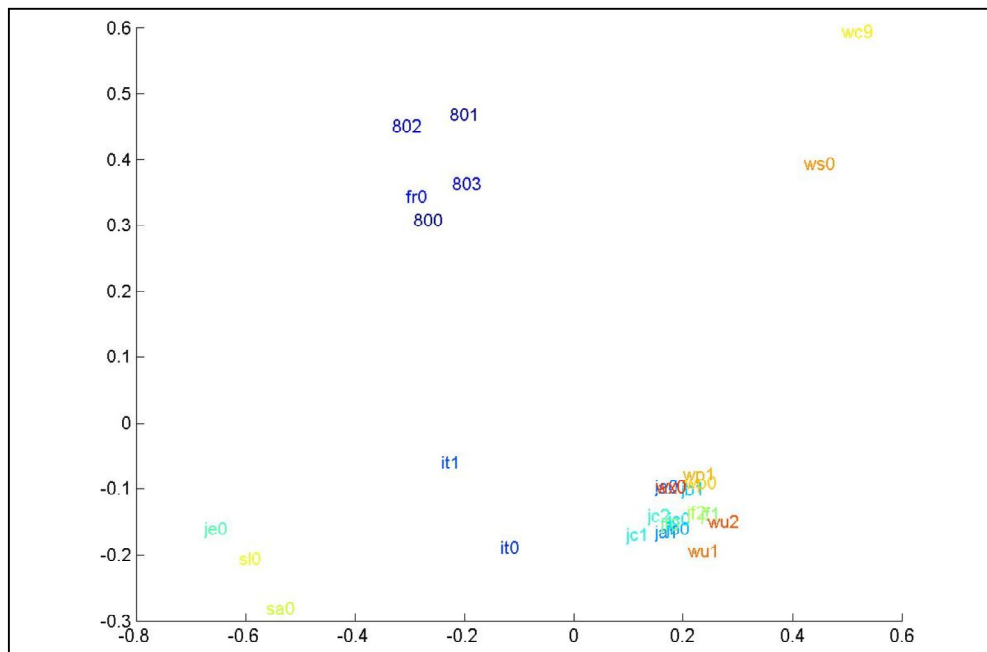


Figura 7. Mapa MDS que representa las distancias entonativas entre distintas variedades románicas realizado con *Amper StatDistances*

Además, mediante unos gráficos de cajas, *Amper StatDistance* ofrece también la posibilidad de valorar la variabilidad observada entre diferentes repeticiones de un

mismo informante y entre las repeticiones de los diferentes informantes de un punto de encuesta. La Figura 8 contiene un ejemplo de gráfico de cajas realizado con *Amper StatDistances* adaptado de Roseano (2012).

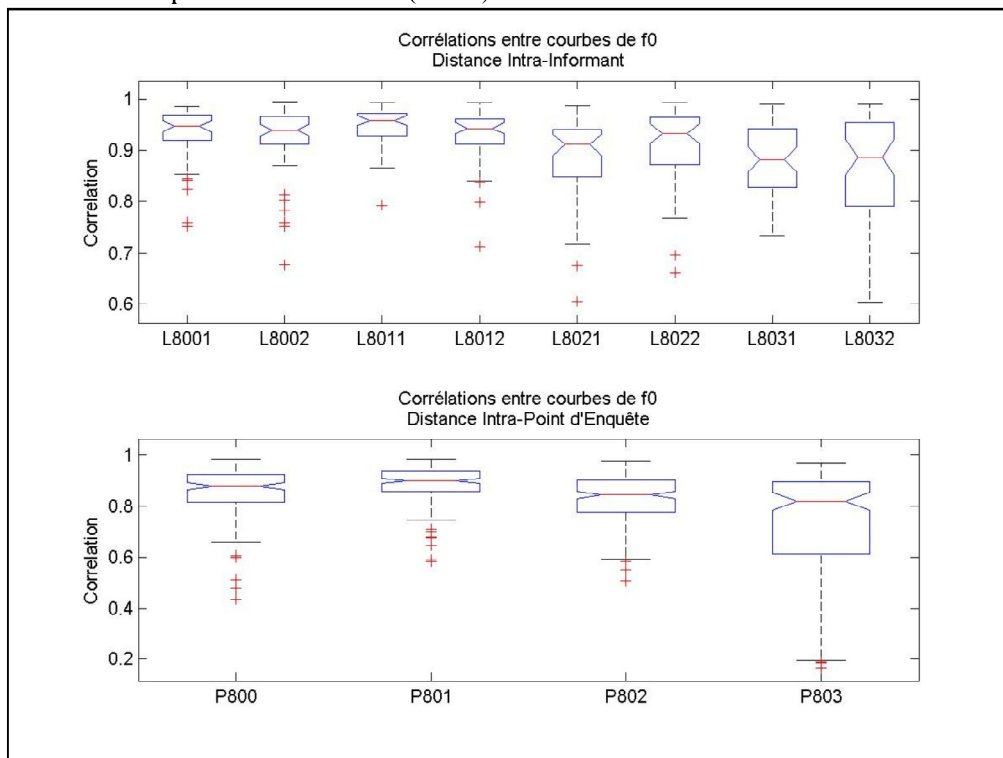


Figura 8. Mapa MDS que representa las distancias entonativas entre distintas variedades románicas realizado con *Amper StatDistance*

La limitación fundamental de *Amper StatDistance* consistía en el hecho de que no permitía visualizar la matriz de distancias ni escoger los métodos de conglomeración. Por esa razón en el Laboratorio de Fonética de la Universitat de Barcelona, centro desde el que se lidera la parte catalana y española del mismo proyecto AMPER, se elaboró – pocos años después de la aparición de *Amper StatDistances*– una rutina capaz de subsanar el inconveniente en cuestión.

La rutina elaborada por el laboratorio de fonética (Roseano, Fernández Planas, Elvira-García y Martínez Celdrán, 2015) toma como base los datos el *output* del análisis numérico que se suele llevar a cabo en el marco AMPER y crea, mediante un script de Praat, una base de datos en un formato Excel que se puede analizar con SPSS. Una vez que los datos numéricos en cuestión se han importado en SPSS, es posible aplicar técnicas de *clustering* equivalentes a las que se implementan en la dialectometría tradicional. La ventaja, en comparación con *Amper StatDistances*, consiste en el hecho de que, así haciendo, es posible seleccionar los algoritmos adecuados y ofrecer los indicadores estadísticos (como el *stress* o el RSQ) que en el programa anterior quedaban ocultos. La rutina del Laboratorio de Fonética, de acuerdo con la propuesta clásica de Hermes (1998a,b), utiliza como índice de la distancia entonativa entre dos frases, que podemos llamar 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 $\{f0x1, f0x2, \dots, f0xN\}$ y $\{f0y1, f0y2, \dots, f0yN\}$, donde N es el número de puntos de medición de F0 en cada una de las dos frases, mientras que $f0xi$ y $f0yi$ son los valores de F0 en st en cada uno esos puntos, la media cuadrática en cuestión viene dada por la siguiente fórmula:

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

Esta fórmula proporciona la distancia entre dos frases. Para determinar la distancia general entre todas las frases de dos puntos de encuesta, la rutina utiliza 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 construye una matriz de distancias que aparece en la Figura 9 y que corresponde a la fase 2 de la Figura 2.

	Siena	Firenze	Arezzo	Perugia	Barcelona	Lleida	Tortosa	Fraga	Palma	Maó	St.Josep	Andorra	l'Alguer	València	Castelló	Alacant	Girona	Tarragona	Ita.Cerd.	Sardo	
Siena	0.79362																				
Firenze	2.01483	0.52805																			
Arezzo	1.91771	1.32969	0.45910																		
Perugia	1.84376	1.40228	1.23804	0.57892																	
Barcelona	2.70149	2.74551	2.20641	2.47984	1.27458																
Lleida	2.57542	2.69118	2.25903	2.43303	2.60293	2.09842															
Tortosa	3.41718	3.17497	2.60438	2.70209	2.56334	2.89857	1.68506														
Fraga	3.62141	3.36905	3.01657	2.87911	2.51865	2.95265	3.04267	1.71903													
Palma	2.39133	2.18763	1.69671	1.62773	2.29823	2.34280	2.16346	2.66916	0.92233												
Maó	2.29306	2.09699	1.55497	1.67030	2.40566	2.26436	2.67430	2.74954	1.59460	1.12967											
St.Josep	2.62173	2.49825	2.08536	2.01748	2.19261	2.22254	3.73926	2.29165	1.63688	1.56400	1.19501										
Andorra	3.29149	3.35646	2.78820	2.77382	1.98106	2.57778	2.56256	2.43157	2.38082	2.47538	2.07799	1.81683									
l'Alguer	3.45937	2.45455	2.78842	2.90488	3.93404	3.53782	3.79562	4.30739	2.95567	2.97357	3.30629	3.99666	1.81205								
València	2.61536	2.29325	1.99987	1.69015	1.81788	2.20153	1.84788	1.94496	1.47872	1.67567	1.45762	1.92488	3.43892	1.43678							
Castelló	2.90009	2.29976	1.98906	1.81123	2.03270	2.10225	2.10483	2.08824	1.43910	1.72195	1.59209	2.13889	3.41749	1.43678	1.17111						
Alacant	2.82157	2.24651	1.78981	1.67390	2.09960	2.25883	2.11356	2.23795	1.44771	1.85988	1.62390	2.13361	3.17047	1.41619	1.40801	1.05075					
Girona	2.67593	3.03603	2.28614	2.39073	2.12698	2.61069	2.56769	2.70127	2.19328	2.22967	1.94563	2.21256	3.78602	1.80547	1.94542	2.05511	1.76715				
Tarragona	3.11204	3.10449	2.37928	2.60110	1.91436	2.54337	2.61550	2.62002	2.48882	2.27695	1.90090	1.99143	3.41376	1.53402	1.88835	1.94739	2.11682	1.16163			
Ita.Cerd.	3.33469	2.42547	2.71852	2.95272	3.76104	3.73926	3.93832	4.36039	3.19864	3.01855	3.38387	4.14699	2.30358	3.40952	3.44873	3.17457	3.82572	3.45911	1.20007		
Sardo	3.14025	2.36387	2.67281	2.66999	3.58960	3.47936	3.62814	4.02838	3.03382	2.62801	3.09136	3.82080	2.33841	3.27006	3.18927	2.96695	3.59627	3.31371	1.86951	1.14617	

Figura 9. Matriz de distancias

La matriz de distancias constituye a su vez la base para la fase final del proceso de análisis, la tercera, que se efectúa con SPSS y consiste en un análisis de clúster cuya finalidad es clasificar los puntos de encuesta en grupos a partir de la diferencia entre sus características entonativas, que es el objetivo último de un análisis dialectométrico entonativo (en decir, corresponde a la fase 3 de la Figura 2). La Figura 10 contiene un ejemplo (adaptado de Roseano et ál., 2015) de representación gráfica, en forma de dendrograma, de los resultados del análisis dialectométrico. Cabe destacar que para el gráfico en cuestión se ha podido escoger y explicitar el método de conglomeración (en este caso, el valor más lejano) y la medida (intervalo distancia euclidiana; valores no transformados), cosa *Amper StatDistances* que no permitía.

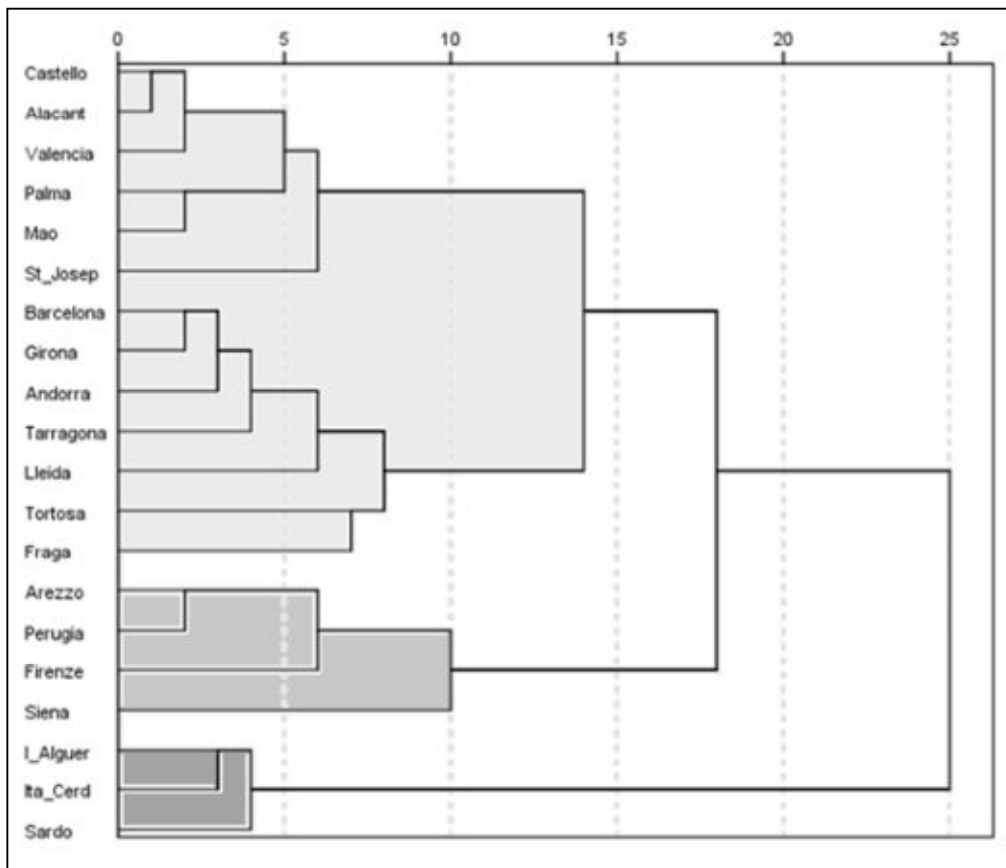


Figura 10. Dendrograma del análisis de clúster de las distancias entonativas entre variedades románicas (adaptado de Roseano et ál., 2015)

A pesar de cumplir con el requisito de explicitar y flexibilizar las tres fases imprescindibles de una dialectometría entonativa que se han ilustrado en la Figura 2, la rutina del Laboratorio de Fonética que se acaba de presentar seguía presentando algunos defectos. El más importante de ellos es, sin duda, su laboriosidad. De hecho, la realización de un análisis dialectométrico con esos instrumentos implica una serie de operaciones informáticas (que en este trabajo no se detallan por razones de espacio) que representan un coste muy alto en términos de tiempo. Por esta razón, tal y como se verá más adelante, en los años siguientes surgirán rutinas más sencillas.

6. La dialectometría entonativa alfabética: su origen

En un segundo momento, aproximadamente hacia el año 2012, en el marco del proyecto *Atles interactiu de l'entonació del català* (Prieto y Cabré, 2007-2012) del *Grup d'Estudis de Prosòdia* de la Universitat Pompeu Fabra, se lleva a cabo el primer análisis dialectométrico de datos entonativos a partir de datos alfabéticos (en este caso, transcripciones fonológicas de configuraciones nucleares de diferentes modalidades oracionales). El instrumento utilizado es *Gabmap* (Nerbonne et ál., 2011), un instrumento capaz de procesar datos de tipo alfabético¹³. Los resultados del análisis en cuestión se publicaron en forma de mapa en Prieto y Cabré (2013) y se reproducen en la Figura 11. De ella, cabe destacar una conclusión a la que llegan los autores del trabajo

¹³Debido al hecho de que *Gabmap* no acepta símbolos comúnmente utilizados en la transcripción de la entonación (como * + - % > ! i), se han tenido que convertir en símbolos alfabéticos convencionales. Por ejemplo: L+H* → LpHs, L% → Lb, etc.

en cuestión y que parece ser una constante de todos los trabajos de dialectometría entonativa: que los grupos dialectales que se pueden formar a partir de la entonación no coinciden con los que propone la dialectología tradicional. En el caso del catalán, por ejemplo, los grupos dialectales entonativos no siguen la tradicional división entre variedades orientales y occidentales de la lengua, que en el mapa de la Figura 11 se representa con una línea de rayas.

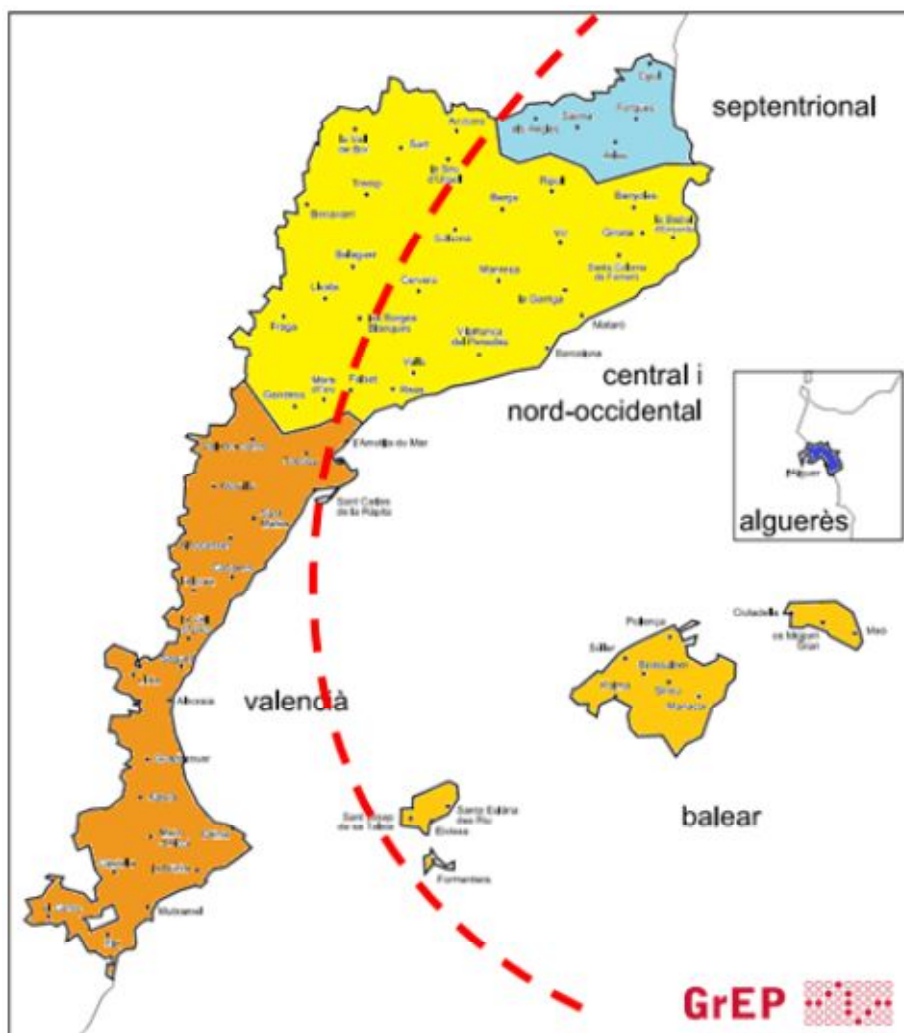


Figura 11. Representación en forma de mapa de una dialectometrización de datos entonativos del catalán (adaptada de Prieto y Cabré, 2013)

7. La dialectometría entonativa alfabética y numérica: primeros intentos

Si en los años entre el 2008 y el 2012 las dos líneas de análisis dialectométrico entonativo (la numérica y la alfabética) se desarrollaban por separado, a partir del 2012 se observan los primeros intentos de compaginar los dos métodos en cuestión. En la tesis doctoral de Roseano (2012) sobre la entonación del friulano se empiezan a utilizar, a la vez, datos numéricos acústicos y datos alfabéticos fonológicos, que se someten a análisis dialectométrico con instrumentos diferentes. Los resultados se comparan entre sí y con los resultados de análisis dialectométricos tradicionales (es decir, de datos no entonativos).

En esta fase, que es la que se caracteriza en verde en la Figura 3, se utilizan paquetes de instrumentos diferentes para analizar, por separado, los datos acústicos numéricos y los datos fonológicos alfabéticos. El primero de ellos, que se utilizó en Roseano (2012), utiliza dos instrumentos que ya existían: *Amper StatDistances* y *Gabmap*. El segundo conjunto de instrumentos es el que se emplea en Fernández Planas, Dorta, Roseano, Díaz, Elvira-García, Martín Gómez y Martínez Celdrán (2015) y que combina un software existente (*DiaTech*) con una rutina creada *ah hoc* en el Laboratorio de Fonética de la UB y que se denomina *Calcu-Dista*. Mientras que, en el marco del estudio en cuestión, *DiaTech* se utiliza para llevar a cabo el análisis de los datos alfabéticos (es decir, de las transcripciones fonéticas realizadas con *AmpeEti*), *Calcu-Dista* se ocupa de dialectometrar los datos numéricos y es un conjunto de scripts de Praat (Elvira-García, 2014) que predisponen una matriz de distancias que se analiza mediante SPSS. En este sentido, *Calcu-Dista* representa la evolución de las rutinas del mismo tipo que se habían creado en el mismo laboratorio unos años antes y que se describen en la Sección 6 de este trabajo. No sorprende, por lo tanto, descubrir que *Calcu-Dista* utiliza el cálculo de distancias prosódicas de Hermes (1998a,b) y la fórmula de RMS para crear una matriz como la que se propone en la Figura 12, con el objetivo de proporcionar representaciones gráficas e indicadores estadísticos como los que aparecen en las Figuras 13 y 14, todas ellas extraídas de Fernández Planas et ál. (2015).

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

Figura 12. Matriz de distancias creada a partir de datos entonativos numéricos (Fernández Planas et ál., 2015)

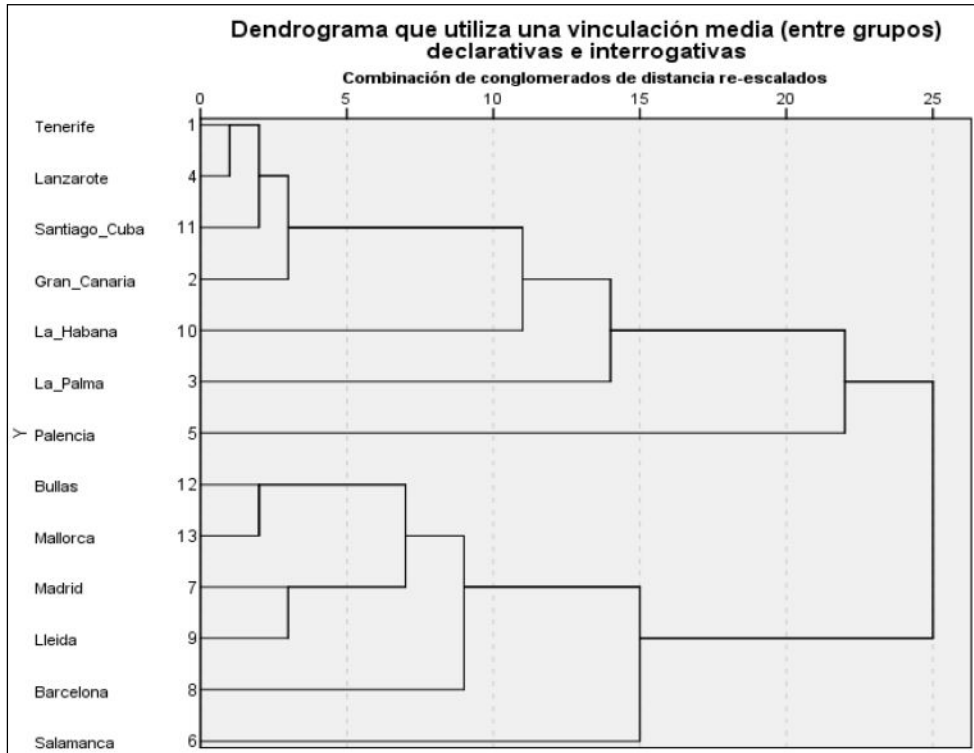


Figura 13. Dendrograma (con indicación de los algoritmos utilizados) creado en SPSS a partir de datos entonativos numéricos (Fernández Planas et ál., 2015)

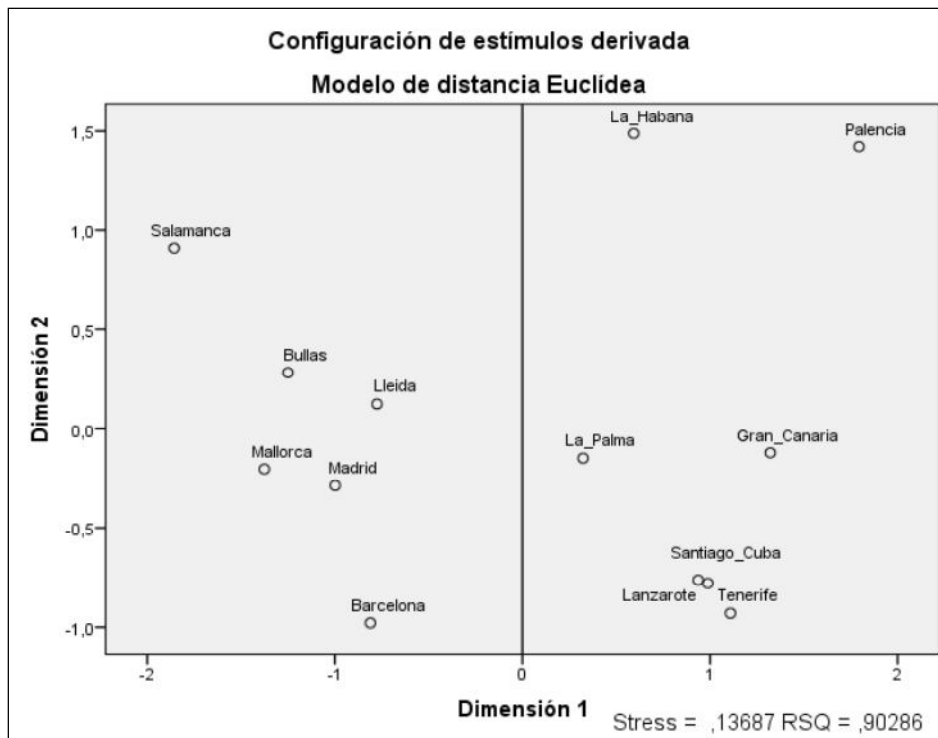


Figura 13. Mapa MDS (con indicación de *stress* y *RSQ*) creado en SPSS a partir de datos entonativos numéricos (Fernández Planas et ál., 2015)

En comparación con las rutinas anteriores del mismo tipo creadas por el equipo del Laboratorio de Fonética de la Universitat de Barcelona, *Calcu-Dista* presenta la ventaja de agilizar ligeramente los procedimientos. Aun así, sigue siendo un sistema poco práctico, en el sentido de que, para realizar una dialectometrización de datos

entonativos acústicos, sigue siendo necesario implementar una serie de acciones con programas diferentes (Praat y SPSS). Además, si bien es cierto que con *Calcu-Dista* se obtiene una representación gráfica de los resultados, cabe decir que esa representación es relativamente básica y que, sobre todo, falta uno de los elementos tradicionales de las representaciones dialectométricas: la trasposición en un mapa geográfico de los resultados del análisis de clúster. Antes de explicar cómo, en los últimos años, se ha intentado superar esas limitaciones (Sección 10), es oportuno presentar brevemente los primeros resultados de una línea de investigación dialectométrica diferente (Sección 9).

8. La dialectometría entonativa perceptiva: primeras aproximaciones

Los estudios dialectales se centran a menudo en los datos de producción, es decir en las diferencias y similitudes que existen entre la manera de realizar un elemento léxico o entonativo en distintos puntos de encuesta. Sin embargo, en las últimas décadas también se ha planteado la cuestión de la percepción de las distancias dialectales (v. Long, 1999; Montgomery y Beal, 2011, entre otros).

En el sector de los estudios suprasegmentales, los intentos de aplicar técnicas dialectométricas a datos perceptivos son escasos. Entre ellos, cabe destacar el de Fernández Planas, Roseano, Elvira-García, Carrera Sabaté y Román Montes de Oca (en prensa), que aplica técnicas de ese tipo a los resultados de diferentes test de percepción de contornos entonativos de lenguas románicas. Desde el punto de vista metodológico, para los experimentos en cuestión se utilizó Praat para la suministración de los test de percepción y la creación de unas matrices de confusiones (como la que se reproduce en la Figura 14, adaptada de Fernández Planas et ál. en prensa) y, a continuación, SPSS para la elaboración en clave dialectométrica en forma de dendrogramas o mapas MDS (como el que se propone en la Figura 15, adaptado de la misma fuente. En ese sentido, el método es comparable al que se ha descrito para *Calcu-Dista*.

	Barcelona	L'Alguer	PortuTurre	Biddanoa	Siena
Barcelona	92,325	50,8125	34,3	11,3125	60,875
L'Alguer	60,0625	91,1375	45,9625	16,95	35,8875
PortuTurre	40,7125	41,1375	92,6375	33,322222	43,3
Biddanoa	32,675	36,75	59,25	95,575	25,8
Siena	61,2875	39,1125	41,5375	15,725	92,3375

Figura 14. Matriz de confusiones creado a partir de datos de percepción obtenidos con Praat (Fernández Planas et ál. en prensa)

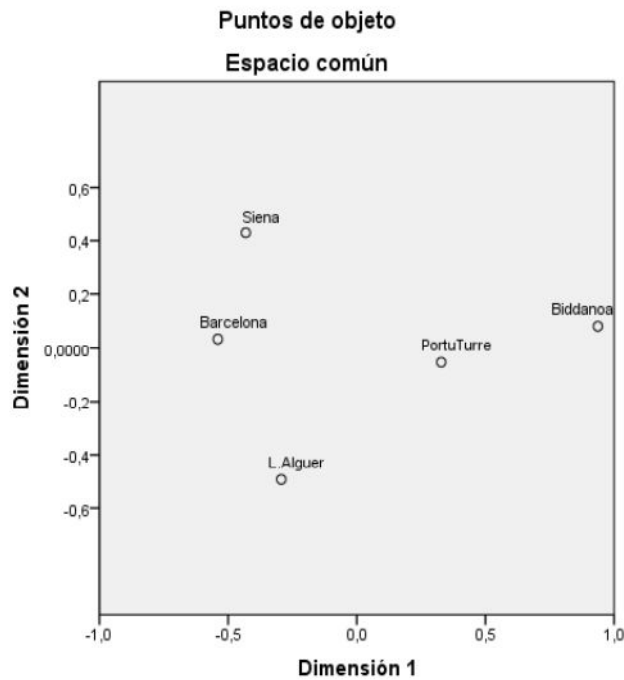


Figura 15. Mapa MDS creado en SPSS a partir de datos de percepción (Fernández Planas et ál. en prensa)

9. La dialectometría entonativa numérica: su desarrollo

En los apartados anteriores se ha visto como los años entre 2008 (cuando se publicaron los primeros resultados obtenidos con *Amper StatDistances*) y 2015 (cuando aparece *Calcu-Dista*) han conocido un desarrollo bastante rápido de la dialectometría entonativa. El 2015 es, justamente, el año que seguramente marca un punto de inflexión en este camino, ya que se presentan oficialmente dos instrumentos dialectométricos elaborados de forma independiente en dos equipos relacionados con el proyecto AMPER: la *Ferramenta* de la Universidad de Santiago de Compostela (Martínez Calvo y Fernández Rei, 2015) y *ProDis* (Elvira-García, Balocco, Fernández Planas, Roseano y Martínez Celdran, 2015), el nuevo aplicativo dialectométrico del Laboratorio de Fonética de la Universitat de Barcelona.

A pesar de presentar diferencias interesantes, los dos instrumentos responden a unas exigencias comunes y presentan unas soluciones comparables. Por un lado, de hecho, se proponen agilizar el análisis dialectométrico de los datos entonativos contenidos en la base de datos de AMPER proponiendo un único *software* que permita llevar a cabo la totalidad de las operaciones que, en fases anteriores, requerían programas distintos (v. Secciones 6 y 8). Por otra parte, amplían el abanico de resultados estadísticos que se podían obtener con SPSS y añaden la posibilidad de representar los resultados mismos en mapas geográficos. Además, ambas se dirigen a un público que no tiene conocimientos de programación informática.

9.1. Ferramenta

La *Ferramenta* presentada por Martínez Calvo y Fernández Rei (2015) es una herramienta informática desarrollada con el *software* estadístico R que permite al usuario realizar un análisis dialectométrico del corpus AMPER. El análisis realizado se fundamenta en las fórmulas de la distancia prosódica propuestas por Hermes (1998a,b)

en los años noventa y que se habían utilizado también en *Amper StatDistances* y *Calcu-Dista*. A partir de allí, calcula la variabilidad observada entre los diferentes informantes de un punto de encuesta y la representa con un gráfico de cajas, al igual que *Amper StatDistances*. La Figura 16 ofrece un ejemplo de ello, adaptado de Martínez Calvo y Fernández Rei (2015).

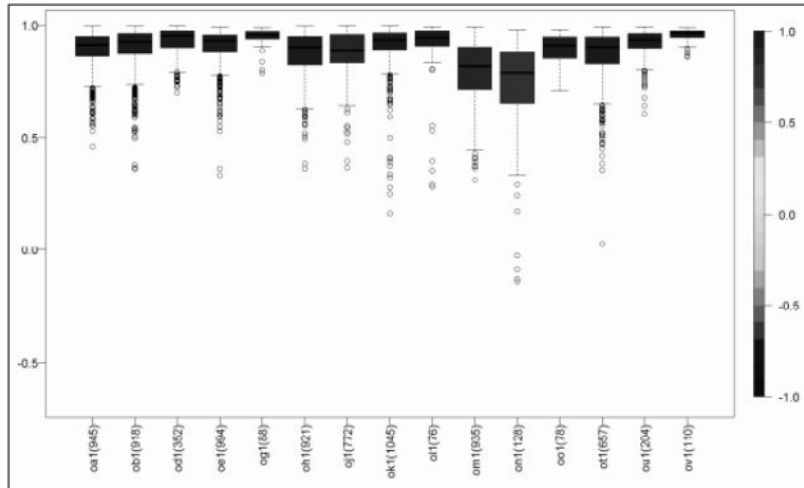


Figura 16. Representación gráfica de la variabilidad en los diferentes puntos de encuesta (adaptado de Martínez Calvo y Fernández Rei, 2015)

A continuación, construye una matriz de correlaciones entre puntos de encuesta y la representa gráficamente de una manera que es mucho más transparente e intuitiva: en vez de generar una matriz numérica, el programa propone una matriz con colores, en la que las correlaciones se indican con gradaciones diferentes que van del rojo al azul. La Figura 17, extraída de Fernández Rei, Moutinho y Coimbra (2014) contiene un ejemplo de la representación en cuestión de la matriz de correlaciones.

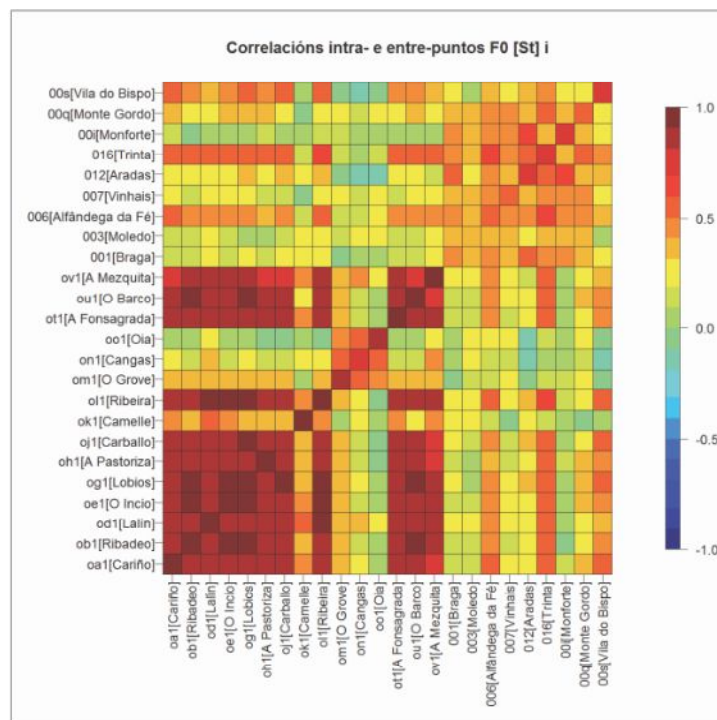


Figura 17. Representación gráfica de las correlaciones entre puntos de encuesta obtenida con *Ferramenta* (adaptado de Fernández Rei, Moutinho y Coimbra, 2014)

A partir de allí, y mediante la aplicación de unos algoritmos (ilustrados en Martínez Calvo y Fernández Rei, 2015), la *Ferramenta* pasa a proporcionar las representaciones gráficas habituales en dialectometría, que son fundamentalmente tres: dendrograma (Figura 18), mapa MDS (Figura 19) y mapa geográfico (Figura 20).

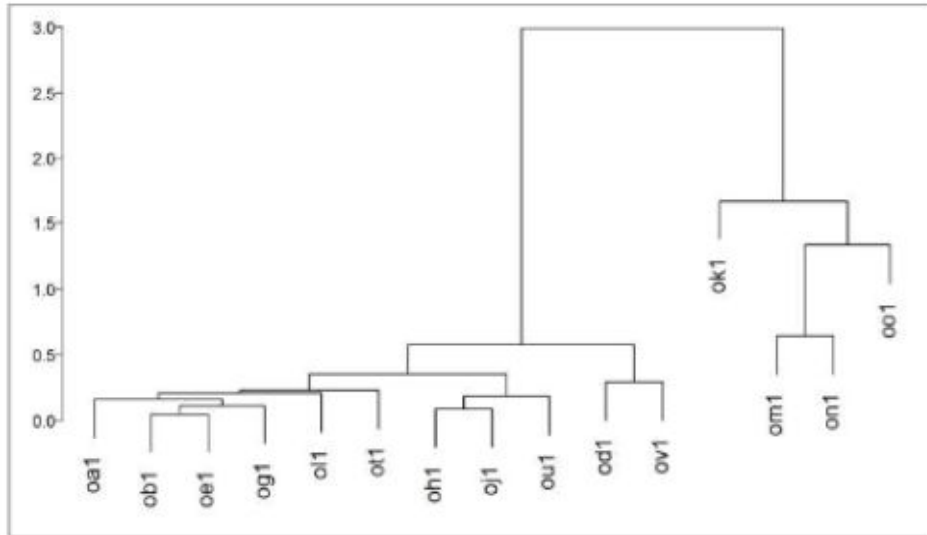


Figura 18. Representación gráfica de los resultados de un análisis dialectométrico con *Ferramenta* en forma de dendrograma obtenido con *Ferramenta* (adaptado de Martínez Calvo y Fernández Rei 2015)

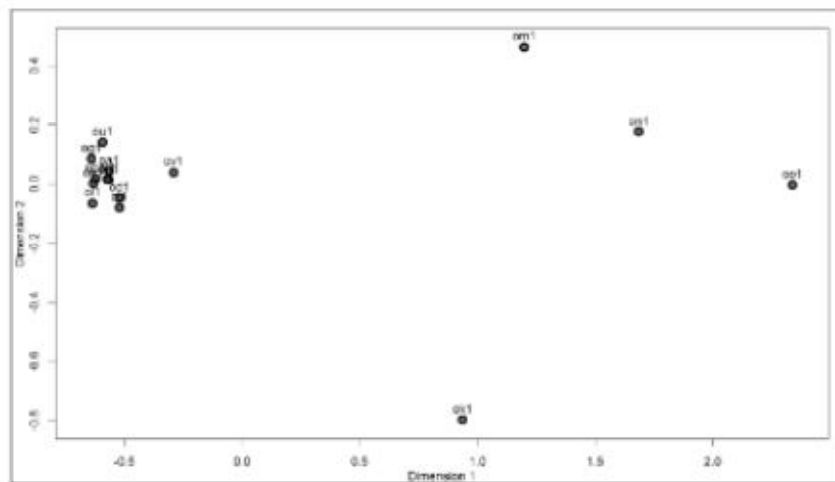


Figura 19. Representación gráfica de los resultados de un análisis dialectométrico con *Ferramenta* en forma de mapa MDS obtenido con *Ferramenta* (adaptado de Martínez Calvo y Fernández Rei 2015)

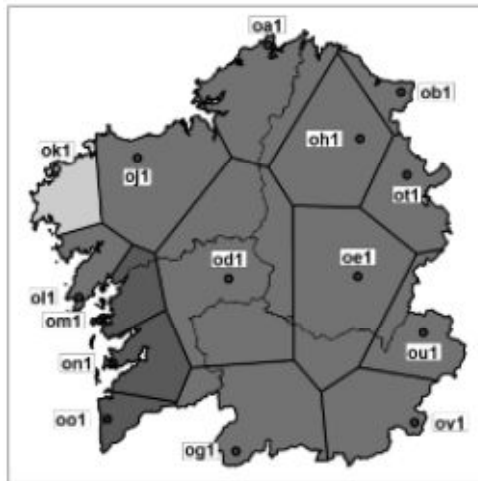


Figura 20. Representación gráfica de los resultados de un análisis dialectométrico con *Ferramenta* en forma de mapa geográfico (adaptado de Martínez Calvo y Fernández Rei 2015)

9.2. *ProDis*

En 2015 el Laboratorio de Fonética de la Universitat de Barcelona, gracias a la cooperación con la Facultad de Matemáticas de la misma universidad, anuncia la creación de *ProDis*, un nuevo programa, realizado en MatLab, para el análisis de datos prosódicos numéricos del proyecto AMPER. *ProDis*, al igual que sus predecesores, calcula la distancia prosódica entre distintos puntos de encuesta utilizando las fórmulas de Hermes (1998a,b). Además, ofrece la posibilidad de ponderar los datos de F0 (el principal parámetro prosódico), por otros parámetros acústicos, como la intensidad y la duración¹⁴.

ProDis, tal y como sus antecesores, calcula la variabilidad entonativa (es decir, si y cuánto los informantes utilizan siempre el mismo patrón entonativo) en los diferentes puntos de encuesta. La diferencia con los programas más antiguos reside en el hecho de que *ProDis* presenta la información en cuestión de forma gráfica, es decir utilizando una matriz en escala de colores en las que los colores fríos indican una alta congruencia de los patrones entonativos y los colores cálidos la alternancia entre patrones entonativos diferentes en el mismo sujeto o punto de encuesta (para más detalles ver Elvira-García et ál., 2015). La Figura 21 ofrece un ejemplo de la representación gráfica en cuestión.

¹⁴Mientras que la ponderación por la intensidad se contemplaba también en los trabajos de Hermes y en la *Ferramenta*, la posibilidad de ponderar por la duración es una de las innovaciones de *ProDis*.

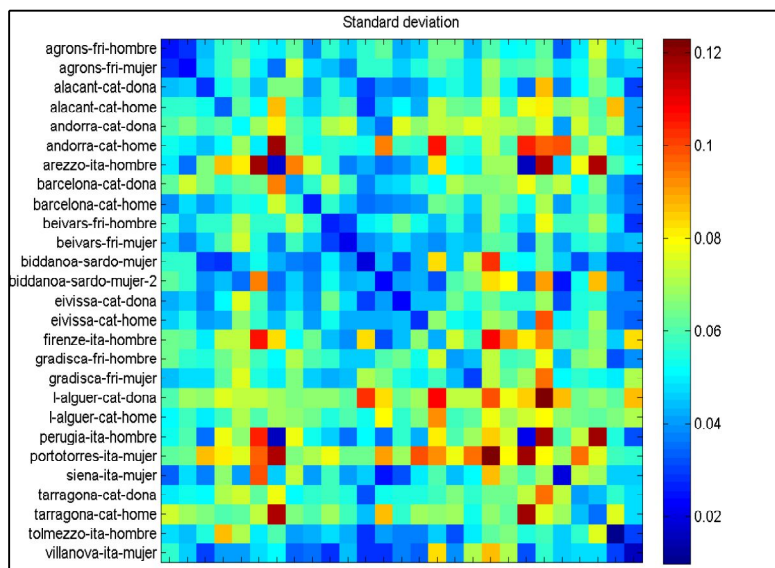


Figura 21. Representación gráfica de la variabilidad en diferentes sujetos y entre diferentes sujetos con *ProDis*

El programa, al igual que la *Ferramenta* descrita en la Sección 10.1, prefiere evitar presentar la matriz de distancias de forma numérica (aunque es una opción que el usuario puede escoger) y utiliza un sistema gráfico basado en gradaciones de colores que van del blanco al rojo oscuro, donde el blanco representa la correlación máxima y el marrón oscuro la más baja. La figura 22 contiene un ejemplo de matriz de correlaciones calculada con *ProDis*.

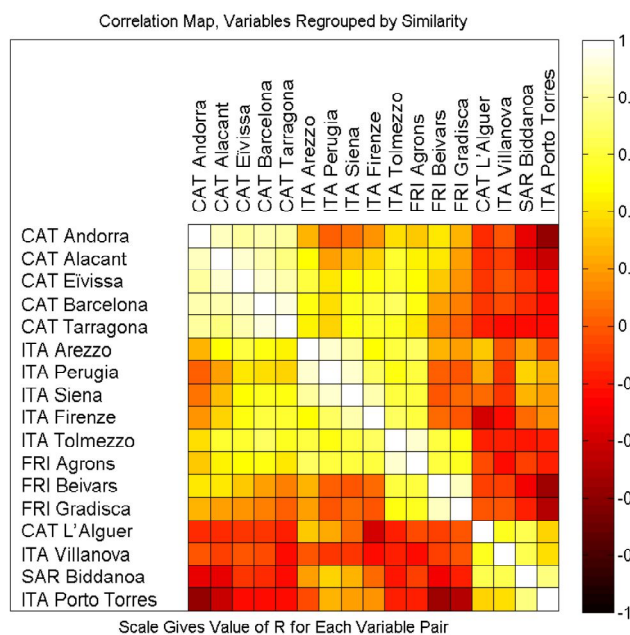


Figura 22. Representación gráfica de las correlaciones entre puntos de encuesta calculads con *ProDis*

A partir de la matriz de correlaciones, y gracias a unos algoritmos que se detallan en Elvira-García et ál. (2015), *ProDis* genera las representaciones gráficas típicas de la dialectometría: dendrograma (Figura 23), mapa MDS (Figura 24) y mapa geográfico (Figura 25).

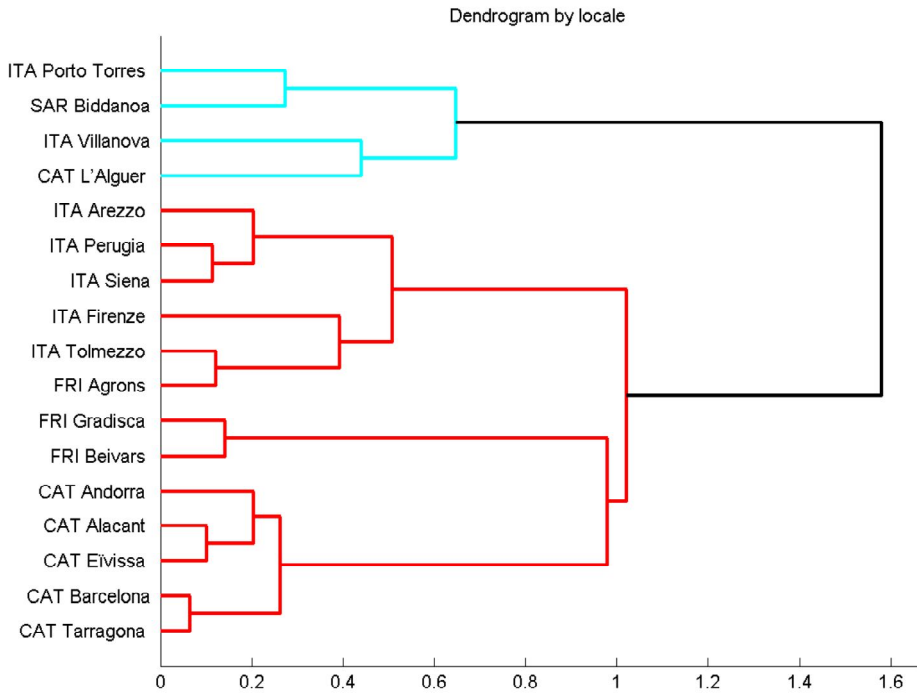


Figura 23. Representación gráfica de los resultados de un análisis dialectométrico con *ProDis* en forma de dendrograma

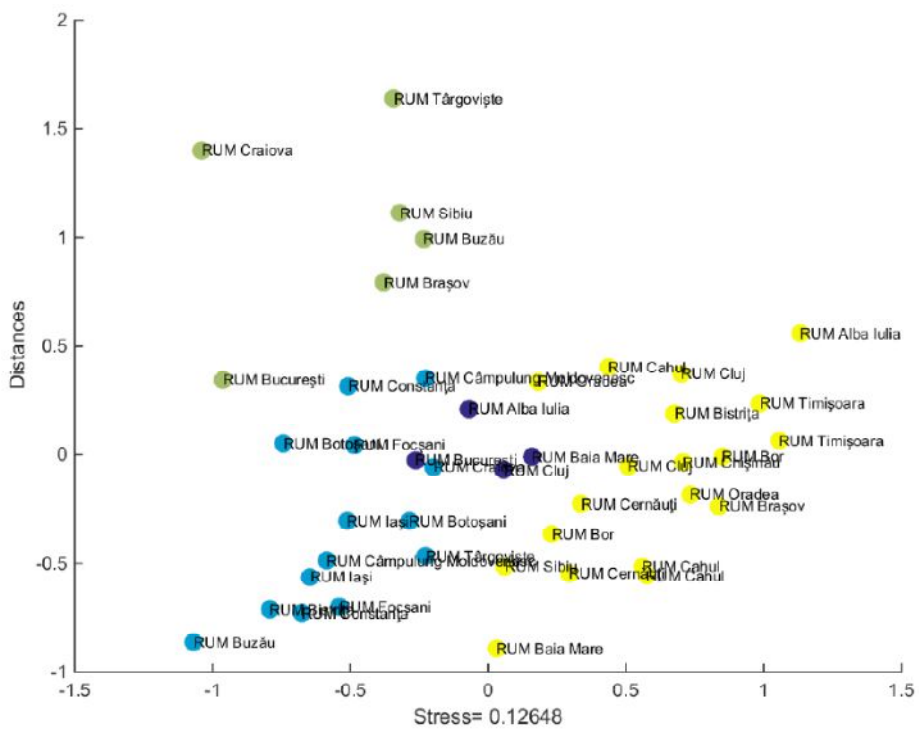


Figura 24. Representación gráfica de los resultados de un análisis dialectométrico con *ProDis* en forma de mapa MDS

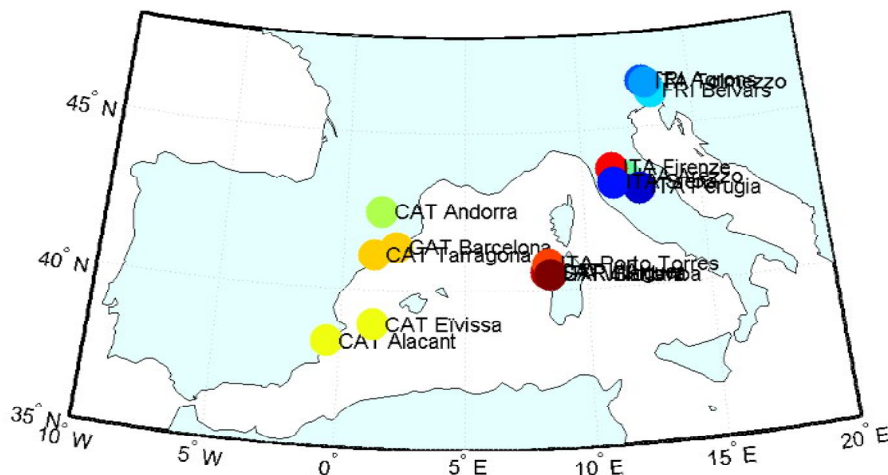


Figura 25. Representación gráfica de los resultados de un análisis dialectométrico con *ProDis* en forma de mapa geográfico

Finalmente, cabe destacar que actualmente (agosto de 2016) *ProDis* se encuentra en fase de ampliación, con el objetivo de añadir otros tipos de análisis estadísticos y *output* gráficos. El objetivo más a largo plazo es la realización de un programa integrado de análisis dialectométrico en el sector prosódico (*ProDis2*) que pueda dialectometrizar, al mismo tiempo, datos entonativos numéricos y alfabéticos, y – en un futuro más lejano – también datos perceptivos.

10. Conclusiones

La dialectometría entonativa es, sin duda, una disciplina muy reciente. Aun así, en pocos años (2008-2015) hemos pasado de no tener ningún instrumento para la dialectometría entonativa a disponer de varios programas, lo que demuestra el interés que suscita el tema en cuestión. Sin embargo, a pesar de los avances muy rápidos, incluso los programas de dialectometría entonativa más avanzados (como *Ferramenta* y *ProDis*) no llegan a tener toda la madurez que caracteriza los aplicativos que se utilizan en dialectometría no entonativa (como, por ejemplo, *DiaTech*).

El primer desafío que se tendrá que superar es de carácter técnico y es la portabilidad de los programas en cuestión, en el sentido de que actualmente ambos se pueden aplicar solo a datos entonativos numéricos que se hayan recogido y analizado de acuerdo con la metodología AMPER. Posibilitar que se puedan analizar con esos instrumentos los datos recogidos con otros métodos representa, sin duda, uno de los retos técnicos más importantes de los próximos años.

El segundo desafío que se tendrá que recoger, más a largo plazo, es el de la integración de los resultados de los dos tipos de dialectometría: la tradicional y la entonativa. Los resultados de la aplicación de los distintos programas de análisis dialectométrico a las bases de datos entonativas de las dos familias (AMPER y IARI) no coinciden, en general, con las clasificaciones dialectométricas de las variedades geográficas a partir de elementos segmentales. En otras palabras, la dialectometría tradicional y la dialectometría entonativa no siempre concuerdan. Esa discrepancia conlleva una serie de dificultades, que se pueden dividir en a) desafíos técnicos y b) retos metodológicos en sentido más amplio. Desde el punto de vista técnico, la dificultad principal parece estar relacionada con la integración, en una misma base de

datos analizable conjuntamente, de datos prosódicos (numéricos) y datos segmentales o léxicos (alfabéticos). Por otra parte, cuando se consiga superar ese reto, se tendrá que establecer qué peso tendrá la componente entonativa en las nuevas bases de datos, pero también otras componentes que no siempre han sido el foco de atención de las grandes bases de datos dialectales, como por ejemplo los aspectos sintácticos.

11. Referencias bibliográficas

- Aurrekoetxea, G., Fernandez-Aguirre, K., Rubio, J., Ruiz, B. y Sánchez, J. (2013). DiaTech: A new tool for dialectology. *Literary and Linguistic Computing*, 28(1), 23-30.
- Clua, E. (2004). El mètode dialectomètric: aplicació de l'anàlisi multivariant a la classificació de les varietats del català. En M.P. Perea, (Ed.), *Dialectologia i recursos informàtics* (pp. 59-88). Barcelona: PPU.
- Contini, M. (1992). Vers une géoprosodie romane. En G. Aurrekoetxea y X. Videgain(Eds.), *Nazioarteko dialektologia biltzarra: Agiriak* (pp. 83-109). Bilbo: Euskaltzaindia.
- Contini, M., Lai, J.P., Romano, A. y Rouillet, S. (2003). Vers un Atlas prosodique des variétés romanes. En J.C. Bouvier, J. Gourc y F. Pic (Eds.), *Sempre los camps auràn segadas resurgantas, Mélanges offerts a Xavier Ravier* (pp. 73-84). Toulouse: CNRS.
- Elvira-García, W., Balocco, S., Fernández Planas, A.M., Roseano, P. y Martínez Celdran, E. (2015). Presentació d'una aplicació informàtica per a l'anàlisi dialectomètrica de dades prosòdiques en el marc de l'Atles Multimèdia de la Prosòdia de l'Espai Romànic. *Presentación en el Workshop sobre la prosòdia del català*. Barcelona, junio, 22.
- Elvira-García, Wendy, 2014, *Calcu-Dista scripts package*. Praat script. <http://stel.uib.edu/labfon/en/praat-scripts> (data de consultación: 29 junio 2016.)
- Fernández Planas, A.M., Dorta, J., Roseano, P., Díaz, X., Elvira-García, W., Martín Gómez, J.A. y Martínez Celdrán, E. (2015). Distancia y proximidad prosódica entre algunas variedades del español: Un estudio dialectométrico a partir de datos acústicos. *Revista de Lingüística Teórica y Aplicada*, 53(2), 13-45.
- Fernández Planas, A.M., Roseano, P., Elvira-García, W., Carrera Sabaté, J. y Román Montes de Oca, D. (en prensa). From a perceptual point of view, is there continuity between languages in contact?. *Spanish in Context*, 14(2).
- Fernández Planas, A.M., Roseano, P., Martínez Celdrán, E. y Romera, L. (2011). Aproximación al análisis dialectométrico de la entonación en algunos puntos del dominio lingüístico catalán. *Estudios de Fonética Experimental*, 20, 141-178.
- Fernández Rei, E., Moutinho de Castro, L. y Coimbra Lúcia, R. (2013). Abordagem dialectométrica das variedades prosódicas do Galego o do Português. *Presentación en las II Jornadas de Ciências da Linguagem*. Aveiro, junio, 12.
- Fernández Rei, E., Moutinho de Castro, L. y Coimbra Lúcia, R. (2014). As entoacións galega e portuguesa: a fronteira á luz da dialectometría e da percepción. En X. Sousa, M. Negro y R. Álvarez (Eds.), *Lingua e identidade na fronteira galego-portuguesa* (pp. 115-141). Santiago de Compostela: Consello da Cultura Galega.
- Fernández Rei, Elisa (Ed.), 2016, *AMPER-Galicia, Atlas Multimedia Prosódico do Espazo Románico*. <http://ilg.usc.es/amper/> (data de consultación: 29 junio 2016.)
- Frota, S. y Prieto, P. (eds.). (2015). *Intonation in Romance*. Oxford: Oxford University Press.

- Frota, Sónia y Marisa Cruz (Eds.), 2012-2015, *Interactive Atlas of the Prosody of Portuguese*. <http://labfon.letras.ulisboa.pt/InAPoP/> (data de consultación: 29 junio 2016.)
- Goebel, H. (1981). Eléments d'analyse dialectométrique (avec application à l'AIS). *Revue de Linguistique Romane*, 45, 349-420.
- Goebel, H. (1982). *Dialektometrie: Prinzipien und Methoden des Einsatzes der numerischen Taxonomie im Bereich der Dialektgeographie*, Wien: Verlag der Öst. Akademie der Wissenschaften.
- Goebel, H. (2013). La dialectometrització dels quatre primers volums de l'ALDC: Una breu presentació. *Estudis Romànics*, 35, 87-116.
- Haimerl, E. (2006). Database design and technical solutions for the management, calculation, and visualization of dialect mass data. *Literary & Linguist Computing*, 21(4), 437-444.
- Hermes, D. J. (1998a). Auditory and visual similarity of pitch contours. *Journal of Speech, Language and Hearing Research*, 41, 63-72.
- Hermes, D. J. (1998b). Measuring the perceptual similarity of pitch contours. *Journal of Speech, Language and Hearing Research*, 41, 73-82.
- Long, D. (1999). Mapping non-linguists' evaluations of Japanese language variation. En D. Preston (Ed.), *Handbook of Perceptual Dialectology* (pp. 199-226). Amsterdam: John Benjamins.
- Martínez Calvo, A., y Fernández Rei, E. (2015). Unha ferramenta informática para a análise dialectométrica da prosodia. *Estudios de Fonética Experimental*, 24, 289-303
- Martínez Celdrán, Eugenio y Ana Ma. Fernández Planas (Eds.), 2003-2016a, *AMPER-CAT. Atlas Multimèdia de la Prosòdia de l'Espai Romànic*. http://stel.ub.edu/labfon/amper/index_ampercat_cat.html (data de consultación: 29 junio 2016.)
- Martínez Celdrán, Eugenio y Ana Ma. Fernández Planas (Eds.), 2003-2016b. *AMPER España e Iberoamérica. Atlas multimedia de la prosodia del espacio románico*. <http://stel.ub.edu/labfon/amper/cast/index.html> (data de consultación: 29 junio 2016.)
- Montgomery, C. y Beal, J. (2011). *Perceptual Dialectology*. En W. Maguire y A. McMahon (Eds.), *Analysing Variation in English* (pp. 121-148). Cambridge, NY: Cambridge University Press.
- Moutinho de Castro, L., Coimbra Lidia, R., Rilliard, A. y Romano, A. (2011). Mesure de la variation prosodique diatopique en portugais européen. *Estudios de Fonética Experimental*, 20, 33-56.
- Moutinho de Castro, Lurdes (Ed.), 2008-2013, *AMPER-POR. Atlas Multimèdia Prosòdico do Espaço Românico - Língua Portuguesa*. <http://pfonetica.web.ua.pt/AMPER-POR.htm> (data de consultación: 29 junio 2016.)
- Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P. y Leinonen, T. (2011). Gabmap: A web application for dialectology. *Dialectologia, Special issue II*, 65-89.
- Perea, M.P. (2010). La dialectometría y su aplicación en el estudio de las variedades dialectales del catalán. *Revista de Filología Asturiana*, 9(1), 109-130.
- Prieto, P. y Cabré, T. (eds.). (2013). *L'entonació dels dialectes catalans*. Barcelona: Publicacions de l'Abadia de Montserrat
- Prieto, Pilar y Paolo Roseano (Eds.), 2009-2013, *Atlas interactivo de la entonación del español*. <http://prosodia.upf.edu/atlasentonacion/> (data de consultación: 29 junio 2016.)
- Prieto, Pilar y Rafèu Sichel-Bazin (Eds.), 2014, *Atlàs interactiu de l'intonacion de l'occitan*. <http://prosodia.upf.edu/atlasintonacion/> (data de consultación: 29 junio 2016.)

- Prieto, Pilar y Teresa Cabré (Eds.), 2007-2012, *Atles interactiu de l'entonació del català*. <http://prosodia.upf.edu/atlesentonacio/> (data de consultación: 29 junio 2016.)
- Prieto, Pilar y Teresa Cabré (Eds.), 2007-2012, *Atles interactiu de l'entonació del català*. <http://prosodia.upf.edu/atlesentonacio/> (data de consultación: 29 junio 2016.)
- Prieto, Pilar, Joan Borràs-Comes y Paolo Roseano (Eds.), 2010-2014, *Interactive Atlas of Romance Intonation*. <http://prosodia.upf.edu/iari/> (data de consultación: 29 junio 2016.)
- Rilliard, A. y Lai, J.P. (2008). La Base de Données AMPER et ses interfaces: Structure et formats de données, exemple d'utilisation pour une analyse comparative de la prosodie de différents parlars romans». En L. Moutinho de Castro y R. Lídia Coimbra (Eds.), *Actas das I Jornadas Científicas AMPERPOR* (pp. 127-139.). Aveiro: Universidade de Aveiro.
- Romano, A. (1999). *Analyse des structures prosodiques des dialectes et de l'italien régional parlés dans le Salento (Italie): Approche linguistique et instrumentale*, Tesis doctoral, Université Stendhal Grenoble 3.
- Romano, A. y Miotti, R. (2008). Distancias prosódicas entre variedades románicas. En A. Turculeț (Ed.), *La variation diatopique de l'intonation dans le domaine roumain et roman* (pp. 231-249). Iași: Editura Universității Alexandru Ioan Cuza.
- Romano, A., Contini, M., Lai, J.P., y Rilliard, A. (2011). Distancias prosódicas entre variedades románicas en el marco del proyecto AMPER. *Revista Internacional de Lingüística Iberoamericana*, 9(1), 13-25.
- Roseano, P. (2012). *La prosòdia del friülà en el marc de l'Atles Multimèdia de Prosòdia de l'Espai Romànic*, Tesis doctoral, Universitat de Barcelona.
- Roseano, P. y Fernández Planas, A.M. (2013). Transcripció fonètica i fonològica de l'entonació: Una proposta d'etiquetatge automàtic. *Estudios de Fonética Experimental*, 22, 275-332.
- Roseano, P., Fernández Planas, A.M., Elvira-García, W. y Martínez Celdrán, E. (2015). Contacto lingüístico y transferencia prosódica: El caso del alguerés. *Dialectologia et Geolinguistica*, 23, 95-123.
- Roseano, Paolo y Fernández Planas, Ana Maria (Eds.), 2009-2013, *Atlant multimediàl de prosodie des varietâts romanichis*. <http://stel.ub.edu/labfon/amper/friul/index.html/> (data de consultación: 29 junio 2016.)
- Seguy, J. (1971). La relation entre la distance spatiale et la distance lexical». *Revue de Linguistique Romane*, 35, 335-357.
- Sullivan, J. N. (2011). *Approaching intonational distance and change*, Tesis doctoral, University of Edinburgh.
- Turculeț, Adrian (Ed.), 2010-2013, *AMPROM. Atlas multimedia prozodic român*. <http://amprom.uaic.ro/indexamprom.php> (data de consultación: 29 junio 2016.)
- Valls, E., Nerbonne, J., Prokic, J., Wieling, M., Clua, E. y Lloret, M.R. (2012). Applying the Levenshtein distance to Catalan dialects: A brief comparison of two dialectometric approaches. *Verba*, 39, 35-61.

Agradecimientos

Proyecto "Tecnologías derivadas de AMPER-CAT y análisis de corpus complementarios" (FFI2015-64859-P MINECO/FEDER), IP Ana Ma. Fernández-Planas.