

# APLICACIÓN DE PRODIS A LA DIALECTOMETRIZACIÓN DE DATOS PROSÓDICOS DEL NORDESTE DE ITALIA

ANA MA. FERNÁNDEZ PLANAS<sup>1</sup>, WENDY ELVIRA-GARCÍA<sup>1</sup> Y PAOLO ROSEANO<sup>1,2</sup>

<sup>1</sup>Universitat de Barcelona, <sup>2</sup>University of South Africa  
anamariafernandez@ub.edu, wendyelvira@ub.edu, paolo.roseano@ub.edu

## ABSTRACT

Este trabajo presenta los resultados del análisis cuantitativo de los datos sobre la entonación de algunas lenguas románicas del Nordeste de Italia que se han recogido en el marco de la sección friulana del Atlas Multimedia de la Prosodia del Espacio Románico (AMPER-Friül). El análisis de los datos se ha llevado a cabo con ProDis, un programa de análisis dialectométrico creado en el Laboratori de Fonètica de la Universitat de Barcelona.

Los resultados permiten destacar que el friulano y el véneto, las dos lenguas románicas propias del área nororiental del Estado italiano, comparten rasgos prosódicos importantes, como, por ejemplo, el perfil de la curva entonativa de las interrogativas absolutas. Por otra parte, las variedades regionales de italiano habladas en Friul y en el Véneto se diferencian claramente, desde el punto de vista prosódico, del italiano de la Toscana y presentan un parecido evidente con las lenguas románicas de sustrato (el friulano y el véneto, respectivamente).

Keywords: prosody, intonation, dialectometry, ProDis, AMPER

## 1. INTRODUCCIÓN

En los años 70 del siglo XX, la preocupación por medir objetivamente las distancias entre variedades lingüísticas llevó al nacimiento de la dialectometría (v. entre otros, Séguy, 1971; Goebel, 1981), que consiste en aplicar técnicas de análisis estadístico a bases de datos de 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, Roullet, 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 poco numerosos los estudios que aplican métodos estadísticos para valorar las diferencias y semejanzas entre la entonación de diferentes dialectos o lenguas (Roseano, 2016). Una de las razones por las que los estudios de este tipo son relativamente escasos ha sido la falta, hasta fechas muy cercanas, de programas informáticos que permitieran llevar a cabo un análisis dialectométrico de datos prosódicos.

Este trabajo presenta los resultados de la aplicación de un programa de ese tipo, ProDis, que ha sido creado en el Laboratori de Fonètica de la Universitat de Barcelona (Fernández-Planas, Roseano, Elvira-García y Balocco, 2017). Mediante ProDis, se han dialectometrizado datos prosódicos de algunas lenguas románicas del Nordeste de Italia que se han recogido en el marco de la sección friulana del Atlas Multimedia de la Prosodia del Espacio Románico (AMPER-Friül) (Roseano y Fernández Planas, 2009-2013).

## 2. METODOLOGÍA

### 2.1. Corpus e informantes

Las frases que se han analizado en este artículo pertenecen al llamado corpus fijo de AMPER (Fernández Planas, 2005), que incluye frases enunciativas de foco ancho y frases interrogativas absolutas neutras con tres repeticiones de cada una de ellas por cada informante analizado. Las frases analizadas para este trabajo contienen palabras trisílabas y tienen tres componentes sintácticos básicos (sujeto, verbo y complemento). El verbo siempre es una palabra llana, mientras que en el resto de posiciones se dan todas las combinaciones acentuales posibles. En total, de cada informante, se han analizado 64 frases: 27 declarativas y 27 interrogativas.

Los puntos de encuesta pertenecen a tres dominios lingüísticos, que, en parte, se solapan. Para la lengua italiana se han analizado datos de las variedades geográficamente centrales de la Toscana y la Umbria (en concreto, de las ciudades de Perugia, Siena y Arezzo), de la variedad hablada en Véneto (en Venecia) y de la variedad hablada en Friul (en la ciudad de Tolmezzo). Para la lengua friulana los datos se recogieron en cuatro puntos de encuesta diferentes: Agrons, en representación del dialecto septentrional, Beivars, en representación del friulano central, Gradisca d'Isonzo, en representación del dialecto oriental, y Tesis, en representación del friulano occidental. Para la lengua véneta, finalmente, se han analizado datos del dialecto de Venecia (en concreto, de la isla de Murano).

En los puntos de encuesta de friulano disponemos de dos hablantes, mientras que –actualmente– en los puntos de encuesta de italiano y de véneto solo disponemos de un informante.

En la Tabla 1 se detalla la composición del corpus por lengua, punto de encuesta y sexo de los informantes. La Figura 1 presenta la distribución geográfica de los puntos de encuesta del Nordeste de Italia.

Tabla 1: Corpus analizado.

Lengua	Punto de encuesta	Informante	Número de frases
Friulano	Agrons	M	64
	Agrons	F	64
	Beivars	M	64
	Beivars	F	64
	Gradisca	M	64
	Gradisca	F	64
	Tesis	M	64
	Tesis	F	64

Véneto	Venecia	M	64
Italiano	Venecia	F	64
	Arezzo	M	64
	Perugia	M	64
	Siena	M	64
	Tolmezzo	M	64
Total			896

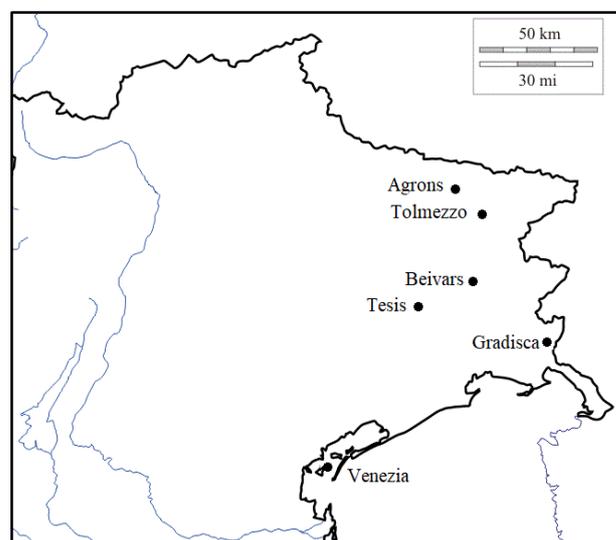


Figura 1: Puntos de encuesta en el Nordeste de Italia.

### 2.2. ProDis

El Laboratorio de Fonética de la Universitat de Barcelona, gracias a la cooperación con la Facultad de Matemáticas de la misma universidad, ha creado ProDis, un programa realizado en MatLab para el análisis de datos prosódicos numéricos del proyecto AMPER. ProDis calcula la correlación de Pearson (1) entre las curvas entonativas de distintos hablantes o distintos puntos de encuesta. Además, ofrece la posibilidad de ponderar los datos de F0 (el principal parámetro prosódico, por otros parámetros acústicos, como la intensidad y la duración.

$$(1) R_{f_1, f_2} = \frac{\sum_i w_i(i) w_0(i) (f_1(i) - m_1) (f_2(i) - m_2)}{\sqrt{\sum_i w_i(i) w_0(i) (f_1(i) - m_1)^2 \sum_i w_i(i) w_0(i) (f_2(i) - m_2)^2}}$$

Una vez calculadas todas las correlaciones, ProDis genera una matriz de correlaciones a partir de la cual, gracias a unos algoritmos que se detallan en Fernández Planas et ál. (2017), genera las representaciones gráficas típicas de la dialectometría: dendrograma, mapa MDS y mapa geográfico. Además, ProDis proporciona también indicaciones acerca de la validez estadística de dichos análisis (los valores de stress en el caso de los mapas MDS y, en el caso del dendrograma, el número óptimo de

clústers obtenido mediante análisis de los gráficos de silueta (Elvira-García, 2017).

### 3. RESULTADOS

Se han llevado a cabo dos análisis, ambos ponderando los datos entonativos por intensidad y duración. En primer lugar, se han dialectometrizado solo los datos de la lengua friulana, con el objetivo de proceder a una clasificación de sus variedades entonativas. En segundo lugar, se han añadido los datos de las demás lenguas, tanto del Nordeste de Italia como de Italia central, para averiguar cómo se relacionan entre sí.

#### 3.1. Datos del friulano

El análisis de los datos del friulano indica que existen dos bloques geoprosódicos (y, de hecho, el análisis de los gráficos de silueta indica que el número óptimo de clústers es dos). Tal y como se puede apreciar en el dendrograma de la Figura 2, uno de los clústers incluye Tesis (friulano occidental) y Agrons (friulano septentrional), mientras que el otro incluye Beivars (friulano central) y Gradisca d'Isonzo (friulano oriental).

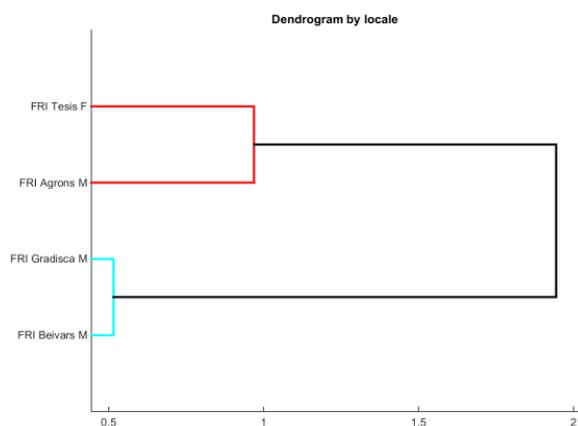


Figura 2: Dendrograma creado a partir de los datos entonativos del friulano ponderados por duración e intensidad.

La agrupación en cuestión se explica a partir de las características de las declarativas del friulano. Mientras en Agrons y Tesis el tonema (o configuración nuclear) acaba con un tono bajo, en Gradisca y Beivars las oraciones aseverativas están caracterizadas por una subida final de F0 (Roseano y Fernández Planas, 2013). Las figuras 3 y 4 proporcionan dos ejemplos de declarativas del friulano (en la Figura 3, una declarativa de Tesis; en la Figura 4, una declarativa de Beivars).

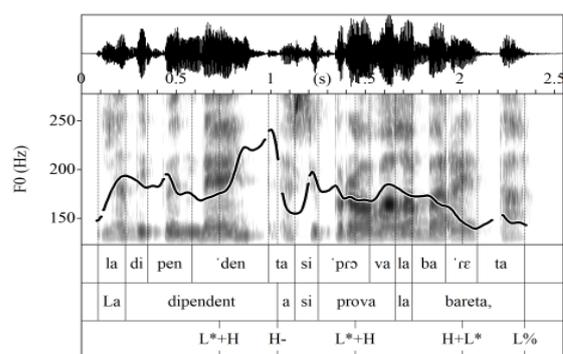


Figura 3: Espectrograma y curva de F0 de una declarativa de friulano de Tesis (friulano occidental).

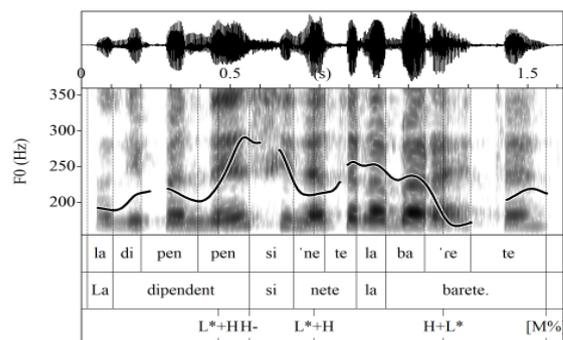


Figura 4: Espectrograma y curva de F0 de una declarativa friulana de Beivars (friulano central).

#### 3.2. Datos del friulano, el véneto y el italiano

El análisis de los datos de las tres lenguas que se incluyen en este estudio permite concluir que el número óptimo de clústers es dos, que son los que aparecen en la Figura 5: por un lado, encontramos los puntos de Italia central (Perugia, Arezzo, Siena) y, por el otro lado, los del Nordeste del Estado.

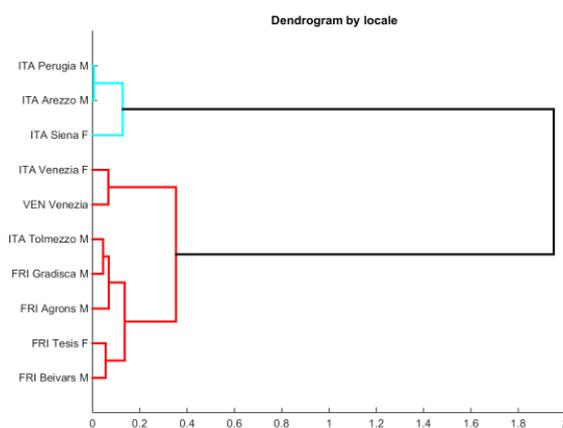


Figura 5: Dendrograma creado a partir de los datos entonativos del friulano, el véneto y el italiano ponderados por duración e intensidad.

A la hora de interpretar esa bipartición, es interesante destacar dos hechos.

En primer lugar, se observa que las variedades regionales de italiano habladas en el Nordeste

de Italia se alejan claramente, desde el punto de vista prosódico, de las variedades de Italia Central, que suelen considerarse las más cercanas a la lengua estándar. Es más, parece bastante claro que la prosodia del italiano regional de Venecia es muy parecida a la del véneto que se habla en la misma localidad, y que la prosodia del italiano regional de Friul (en concreto, de Tolmezzo) se acerca de manera evidente a la del friulano.

En segundo lugar, cabe destacar que el véneto y el friulano, a pesar de ser dos lenguas distintas, acaban formando parte de una misma agrupación geoprósodica (a pesar de que los datos del véneto constituyen una rama claramente separada dentro del bloque). Con toda probabilidad, esto se debe al hecho de que las interrogativas de todas las lenguas y variedades del Nordeste de Italia presentan una curva de F0 claramente diferente del patrón que aparece en Italia Central. En Italia Central la curva está caracterizada por un solo pico en el pretonema y una configuración nuclear que presenta un pequeño pico asociado a la tónica, un descenso y un rápido ascenso final (Figura 6). En el Nordeste, los picos del pretonema son dos y la configuración nuclear presenta una sílaba tónica baja seguida de un ascenso (Figuras 7 y 8).

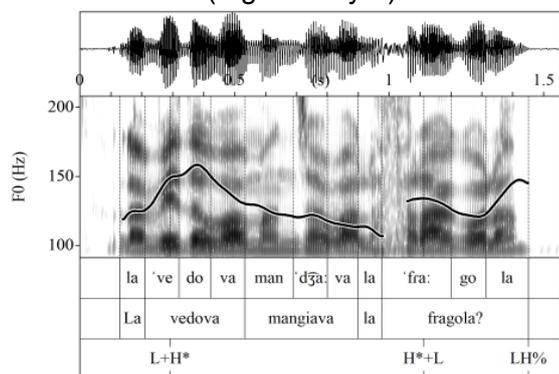


Figura 6: Espectrograma y curva de F0 de una interrogativa de italiano de Perugia.

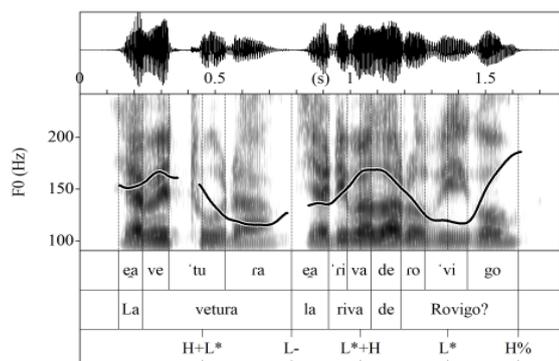


Figura 7: Espectrograma y curva de F0 de una interrogativa de véneto de Venecia.

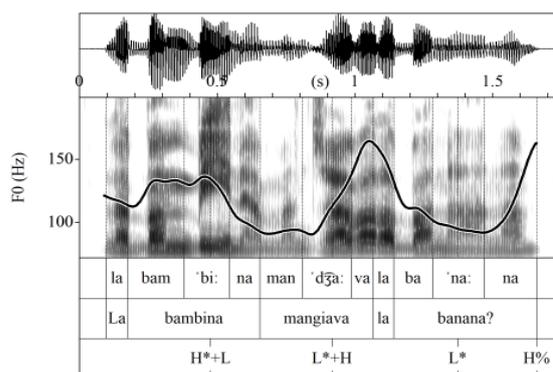


Figura 8: Espectrograma y curva de F0 de una interrogativa de italiano de Tolmezzo.

#### 4. DISCUSIÓN Y CONCLUSIONES

Los resultados del análisis dialectométrico llevado a cabo con ProDis permiten destacar que el friulano y el véneto, las dos lenguas románicas propias del área nororiental del Estado italiano, comparten rasgos prosódicos importantes, como por ejemplo el perfil de la curva entonativa de las interrogativas absolutas. Por otra parte, las variedades regionales de italiano habladas en Friul y en Véneto se diferencian claramente, desde el punto de vista prosódico, del italiano de la Toscana y presentan un parecido evidente con las lenguas románicas de sustrato (el friulano y el véneto, respectivamente).

#### 5. REFERENCIAS BIBLIOGRÁFICAS

- Contini, M., Lai, J.-P., Romano, A., Roulet, S. 2003. Vers un Atlas prosodique des variétés romanes. En J. C. Bouvier, J. Gourc, F. Pic (Eds.), *Sempre los camps auràn segadas resurgantas*, Mélanges offerts a Xavier Ravier (pp. 73-84). Toulouse: CNRS.
- Contini, M. 1992. Vers une géoprosodie romane. En G. Aurrekoetxea, X. Videgain (Eds.), *Nazioarteko dialektologia biltzarra: Agiriak* (pp. 83-109). Bilbo: Euskaltzaindia.
- Elvira-García, W. 2017. L'índex del gràfic de silueta com a eina com a eina metodològica per avaluar l'optimitat dels clústers dialectals. Presentación en el 9è workshop de l'entonació del català. Barcelona, 4 de julio de 2017.
- Fernández Planas, A. M. 2005. Datos generales del proyecto AMPER en España. *Estudios de Fonética Experimental XIV*, 13-27.
- Fernández-Planas, A. M., Roseano, P., Elvira-García, W., Balocco, S. 2017. Génesis y aspectos fundamentales de ProDis. Presentación en el congreso Subsidia: Tools and Resources for Speech Sciences. Málaga, 21-23 junio de 2017.
- Goebel, H. 1981. Eléments d'analyse dialectométrique (avec application à l' AIS). *Revue de Linguistique Romane* 45, 349-420.

- Prieto, P., Borràs-Comes, J., Roseano, P. (Eds.) 2010-2014. Interactive Atlas of Romance Intonation. <http://prosodia.upf.edu/iari/>
- Roseano, P. 2016. Dos décadas de dialectometría entonativa. En A. Iglesias, A. Romero, A. Ensunza (Eds.), *Linguistic Variation in the Basque language and Education II* (pp. 56-80). Bilbao: Universidad del País Vasco.
- Roseano, P., Fernández Planas, A. M. 2013. L'intonazione delle dichiarative neutre e delle interrogative polari in quattro varietà friulane: Agrons, Beivars, Tesis e Gradisca d'Isonzo. *Ladinia* 37: 161-182.
- Roseano, P., Fernández Planas, A. M. (Eds.) 2009-2013. *Atlant multimediâl de prosodie des*

varietâts romanichis.  
<http://stel.ub.edu/labfon/amper/friul/index.html/>  
 Seguy, J. 1971. La dialectométrie dans l'Atlas linguistique de Gascogne. *Revue de Linguistique Romane* 37, 1-4.

#### Agradecimientos

Este estudio ha sido financiado por el Ministerio de Economía y Competitividad con el proyecto FFI2015-64859-P obtenido en convocatoria pública competitiva.