

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Distancias prosódicas entre variedades románicas en el marco del proyecto AMPER

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/98224> since

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:

Questa è la versione pre-print dell'opera:

*“Distancias prosódicas entre variedades románicas en el marco del proyecto
AMPER”.*

*[Romano A., Contini M., Lai J.P. & Rilliard A., Revista Internacional de Lingüística
Iberoamericana, Vol. IX, No. 1 (17), 2011, 17-26 [ISSN 1579-9425]].*

The definitive version is available at:

La versione definitiva è disponibile alla URL:

*[http://www.iber-america.net/cgi-
bin/search.cgi?n=verlagskatalog&lang=en&kat=RILI](http://www.iber-america.net/cgi-bin/search.cgi?n=verlagskatalog&lang=en&kat=RILI)*

Antonio Romano^{*}/Michel Contini^{**}/Jean Pierre Lai^{***}/Albert Rilliard^{****}

Distancias prosódicas entre variedades románicas en el marco del proyecto AMPER¹

Resumen: La entonación en el espacio lingüístico románico presenta esquemas típicos diferentes, marcados diatópicamente y que pueden presentarse convergentes o divergentes entre distintas hablas de la misma área e incluso entre distintos idiomas del mismo hablante (Contini 1992; Romano 2001a, 2001b; Turculeț *et al* 2004; Dorta 2007).

A consecuencia del lanzamiento de un proyecto internacional para la descripción y la comparación de los rasgos prosódicos de las variedades lingüísticas del espacio románico (AMPER), a partir del 2001 varios equipos de distintos laboratorios de Europa y de América han adoptado estrategias comunes de constitución de corpus y los mismos para-digmas de evaluación de los datos.

Si los trabajos en curso no dejan de detallar las peculiaridades locales detectadas en los espacios subregionales, un estudio sobre la variabilidad global que ponga las cosas en su punto sigue siendo un objetivo ambicioso. Intentaremos proponer aquí un marco para la clasificación y la evaluación de los datos ya disponibles con una misma metodología que ya se había utilizado de forma experimental en Romano (2001a) y

* Antonio Romano es responsable del Laboratorio de Fonética Experimental "Arturo Genre" de Turín (LFSAG). Después de sus actividades en Italia, Francia y Reino Unido, en la pública y en la privada, desde el año 2000 es Coordinador Internacional del Proyecto Internacional AMPER (*Atlas Multimedia de Prosodia del Espacio Románico*), junto a Michel Contini. Ha enseñado Lingüística General y Fonética Experimental en diversas universidades italianas y francesas y en la actualidad se dedica especialmente a la Glotología y la Lingüística en la Facultad de Lengua y Literatura Extranjera y en la Facultad de Medicina y Cirugía de la Università degli Studi de Turín.

** Michel Contini es Profesor Emérito de la Université de Grenoble. Es el Director del *Atlas Linguistique Roman*, el Coordinador Internacional, junto con Antonio Romano, del Proyecto AMPER y co-director del proyecto *Atlante Linguistico Multimediale Della Sardegna* (ALiMuS). También es miembro de los Comités Científicos de otros muchos proyectos internacionales de geolingüística, entre ellos el *Atlas Linguarum Europae* y de muchas revistas científicas del dominio románico.

*** Jean-Pierre Lai es Ingeniero de Estudios en la Université de Grenoble. Tras realizar su Tesis Doctoral sobre la entonación del habla de Nuoro (Cerdeña), se ocupa desde 2004 de la Secretaría Científica del Proyecto Internacional AMPER (*Atlas Multimedia de Prosodia del Espacio Románico*). También forma parte del Comité francés del *Atlas Linguistique Roman*. Sus principales trabajos versan sobre dialectología entonativa.

**** Albert Rilliard es investigador del CNRS en el LIMSI (Orsay, Francia) donde dirige los trabajos sobre la percepción y caracterización de la prosodia expresiva en un contexto intercultural: como un locutor transmite diferentes informaciones a través de un mismo canal, en un contexto enunciativo, lingüístico y afectivo dado.

¹ Agradecemos cordialmente a nuestro colega Renzo Miotti (Universidad de Verona) por haber aceptado la gravosa tarea de traducir este artículo al español.

que se aplica ahora a la comparación entre datos *AMPER* (cf. Lai/Rilliard 2008; Miotti/Romano 2008; Rilliard/Lai 2008; Romano/Miotti 2008).

Palabras claves: variedades románicas, evaluación de las distancias prosódicas, formulación, método de evaluación.

Abstract: Intonation in the Romance language area presents different typical patterns, diatopically marked, which in a converging or diverging aspect can occur between different spoken varieties in the same area and even between different languages of the same speaker (Contini 1992; Romano 2001a, 2001b; Turculeț *et al* (ed.) 2004; Dorta 2007).

As a result of the launch of an international project aimed at the description and comparison of the prosodic features of the different language varieties in the Romance space (*AMPER*), since 2001 several teams of different laboratories in Europe and America have adopted common strategies for corpus constitution and the same paradigms for data evaluation.

If work in progress doesn't fail to detail the local features detected in the sub-areas, a study on the overall variability to put things in their point remains an ambitious target. We will propose here a framework for the classification and evaluation of available data with the same methodology that had already been used in an experimental way in Romano (2001a) and is now applied to the comparison between data *AMPER* (cf. Lai/Rilliard 2008; Miotti/Romano 2008; Rilliard/Lai 2008; Romano/Miotti 2008).

Keywords: Romance varieties, evaluation of prosodic distances, formulation, method of evaluation.

1. Introducción

AMPER es un programa científico de geolingüística dialectal cuya finalidad es el estudio de uno de los aspectos menos conocidos del fonetismo de las variedades románicas: la entonación.

El proyecto nació en un centro de dialectología y se enmarca perfectamente en las orientaciones de investigación de la disciplina: su puesta en marcha se justificaba por la existencia de una laguna de la que eran conscientes todos los dialectólogos, es decir, que en sus trabajos, y en especial en los atlas, siempre habían pasado por alto fenómenos prosódicos como la entonación (cf. Contini 1992)².

² La realización de un atlas exige sobre todo el empleo de un mismo cuestionario y de una misma estrategia de encuesta en cada localidad, para obtener la designación local de un determinado referente. Sin lugar a dudas, en el caso de un estudio sobre la variación prosódica, ha sido necesario confeccionar unos cuestionarios especiales, para obtener las frases en las modalidades esperadas, y para adaptarse a las hablas que se quiera estudiar. Por otra parte, cualquiera que sea la variedad, las frases del corpus están basadas sobre versiones diferentes de un mismo *questionnaire commun de base* y por esto presentan la misma estructura sintáctica y unas características fonéticas parecidas para todos los segmentos que las forman. En el *Centro de Dialectología* de Grenoble, en el curso de varias tesis de doctorado y a partir de las propuestas de Lai/Romano/Roullet (1997) (véase también Romano/Contini 2001;

Al igual que los otros proyectos de geolingüística, *AMPER* interesa no solamente a los dialectólogos, que han demostrado su capacidad de innovación desarrollando ciertos sectores de su disciplina, sino que les interesa cada vez más a los fonetistas ‘puros’, no dialectólogos, que vislumbran nuevas posibilidades de colaboración y de apertura hacia nuevos horizontes de investigación en el amplio dominio que ofrece la dialectología: el de las miles de lenguas de tradición oral que son las variedades dialectales y que son de un inmenso interés precisamente en virtud de su estatuto oral, muy alejado de la tendencia a la normalización que prevalece en las lenguas nacionales, donde la influencia de la escritura desempeña un papel mayor (cf. Contini 2008)³.

Para respetar el objetivo propio de un atlas lingüístico y de los métodos de análisis prosódico de los fonetistas se han creado unas pautas comunes de análisis instrumental de la prosodia basadas en la evaluación de distintas variables (F_0 , D e I)⁴: (1) medidas en puntos estratégicos de los enunciados, (2) cuya media se saca sobre la base de varias tomas y (3) evaluadas con unos tests de escucha realizados mediante modelos prototípicos⁵.

Trabajos recientes a partir de *AMPER*, concernientes a comparaciones entre variedades de los varios dominios (véase bibliografía), han llevado a unos resultados muy interesantes y muy alentadores para los objetivos que nos hemos fijado⁶. Por esto,

Contini/Lai/Romano/Roullet/Moutinho/Coimbra/Pereira Bendiha/Secca Ruivo 2002; Lai 2002, Romano 2003, y más recientemente Contini 2008), se ha desarrollado un procedimiento de análisis instrumental de la prosodia y se han añadido unos métodos originales de estilización y modelización que, al mismo tiempo que tienen en cuenta la variación (con una evaluación estadística de la significatividad de ciertos rasgos), permiten también determinar unos andamios medias que caracterizan prototipos entonativos utilizados para tests de percepción.

³ En la mayoría de los casos, los trabajos de fonetistas y fonólogos no se plantean el problema de un análisis contrastivo interlingüístico: las investigaciones relativas a una determinada lengua nacional están basadas en enfoques teóricos e instrumentales diferentes, con las más distintas estrategias relativas a la elección del material de estudio (véase, entre otros, Romano 2003).

⁴ Las rutinas de análisis que se utilizan actualmente (en ambiente MATLAB) son el resultado de la implantación de métodos de medición y de procesamiento de datos prosódicos desarrollados sobre la base del procedimiento de análisis puesto a punto y testado en los estudios de Contini y Boë de los años 70 (para una bibliografía más completa, véase Lai 2005).

⁵ Las informaciones prosódicas extraídas sobre la base de numerosas repeticiones de una muestra de enunciados de medida variable aparecen resumidas en ficheros de texto de pequeño tamaño (.txt) que permiten restituir la información suprasegmental en audio (ficheros .ton) después de la creación de una síntesis por impulsos que utilizamos desde 1997 (del mismo tipo que la que ya está disponible gracias a determinadas funciones de las últimas versiones de *PRAAT*). La realización de estímulos mediante frases de las que no se mantiene sino el esquema prosódico después de la eliminación de todo contenido semántico permite realizar un test de percepción de enorme interés. Sólo se evalúa de este modo el contenido prosódico, con unos experimentos de escucha (del mismo tipo que los realizados por Ohala/Gilbert 1979, véase Romano 1997).

⁶ *AMPER* cuenta con una red internacional que une a dialectólogos y fonetistas de los dominios lingüísticos galo-románico, italiano, rumano, portugués y español (castellano, catalán, gallego y asturiano, de Europa y del Sur de América). Siete comités nacionales son responsables de las encuestas y del análisis de los datos correspondientes a sus áreas de competencia; las áreas dialectales, así como los puntos de encuesta representativos, se definen en base a los conocimientos adquiridos por los dialectólogos. Un comité científico comprende el responsable de cada uno de los comités nacionales y los dos coordinadores internacionales (M. Contini, A. Romano), con un Secretariado Científico (J.P. Lai) y con un responsable de la Base de Datos (A. Rilliard). La producción científica en el ámbito del proyecto se aproxima actualmente a las doscientas publicaciones, entre artículos y comunicaciones, y ha contribuido de modo consistente a la fama del proyecto: prueba de ello es la acogida privilegiada que le

para definir la identidad entonativa de una variedad determinada habrá que relevar los rasgos entonativos pertinentes que la caracterizan con respecto a los de las otras variedades, lo cual representa una primera etapa hacia una tipología prosódica de las variedades dialectales románicas (sobre este tema véase Ohala 1983; Vaissière 1995; Gussenhoven 2002; Vaissière/Boula de Mareüil 2004; Romano 2004).

2. Distancias prosódicas: ejemplos de aplicación

Nuestra propuesta consiste en someter los datos de la base de datos *AMPER* a una evaluación de distancia prosódica, como la que hicimos en Miotti/Romano (2008) y Romano/Miotti (2008) y como propusieron Lai/Rilliard (2008) y Rilliard/Lai (2008) basándose en el método empleado por Hermes (1998a y b)⁷.

Para establecer la distancia entre los datos de dos variedades se comparan las secuencias de valores de F_0 , de duración o de intensidad por medio de una variable oportunamente definida y se evalúa el resultado en relación con un umbral definido, estableciendo el intervalo de oscilación de la variable cuando se trata de repeticiones diferentes de la misma oración, en la misma variedad y por parte del mismo hablante.

La fórmula utilizada para el cálculo del grado de correlación entre dos series X y Y de n datos es la siguiente (cf. Romano/Interlandi 2005):

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y}$$

donde: $-1 \leq \rho_{x,y} \leq 1$ (expresado en porcentaje $-100\% \leq \rho_{x,y} \leq 100\%$) y:

$$Cov(X,Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \text{ (una de las definiciones de la covariancia)}$$

X e Y representan las series de datos de energía o de duración o de F_0 de dos realizaciones (de frases comparables)⁸.

han reservado los Congresos de Fonética Experimental de Santiago de Compostela (2005) y de Granada (2008). Prueba de ello son también los volúmenes monográficos aparecidos: el de su presentación general (*Géolinguistique HS3*, cf. *AMPER* 2005); *La prosodia en el ámbito lingüístico románico* (Dorta 2007), y *La variation diatopique de l'intonation dans le domaine roumain et roman* (Turculeț 2008), así como el volumen 14 de *Estudios de Fonética Experimental* (Martínez Celdrán 2005), dedicado a los trabajos realizados en el cuadro de *AMPER-en-España*; el volumen sobre las investigaciones llevadas a cabo en *AMPER-POR* (Moutinho/Coimbra 2008) y el volumen 2 (Special Issue "Experimental Prosody") de *Language Design* (Pamies/Amorós/Pazos 2008).

⁷ Los ejemplos de aplicación propuestos en estos casos se refieren respectivamente a variedades salentinas, vénetas, ibéricas, friulanas y rumanas, en los trabajos de A. Romano y R. Miotti, sardas y occitanas en los trabajos de J.P. Lai y A. Rilliard.

⁸ Las distancias objetivas propuestas por Hermes (1998a) son las siguientes (cf. Rilliard/Lai 2008):

1) una medida de la correlación entre los vectores de parámetros acústicos extraídos de las dos frases a comparar (correlación de Pearson entre dos vectores x e y de longitud n)

2.1. Distancias prosódicas: aplicación a los enunciados de un mismo hablante distintos por modalidad

En particular, en Romano (2001a) esta fórmula se ha aplicado a todas las repeticiones disponibles para una misma frase y para cada uno de los tres parámetros F_0 , duración e intensidad y todas las medidas de distancia prosódica entre variedades así calculadas se han comparado con un valor medio de distancia entre repeticiones (correlación cruzada o *intercorrelación = medida de desvío global entre las curvas de F_0 de varias realizaciones de enunciados comparables*).

Un ejemplo de la primera aplicación de estas distancias prosódicas a las hablas salentinas queda ilustrado en la Tabla I (cf. Romano 2001a: 226-235), donde por un mismo hablante se comparan los valores de intercorrelación interna, como medida de su coherencia, para 5 tomas de dos frases en los dos códigos en los que ha producido los enunciados (*dialecto e italiano regional*).

Se notará cómo los tres parámetros considerados presentan diferentes condiciones de variabilidad dentro del mismo código con valores medios que es necesario conocer antes de establecer la mayor o menor correspondencia de enunciados producidos por hablantes diferentes.

En particular, comparando la intercorrelación entre los valores de F_0 de repeticiones diferentes de dos hablantes de dialectos cercanos, se han obtenido los valores reproducidos en la tabla II para la modalidad *D* y en la tabla III para la modalidad *I*. La intercorrelación ha resultado ser 78,3% en el caso de *D* (en comparación con el 92,3%

$$r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

2) una medida de la distancia de los promedios cuadrados entre estos dos mismos vectores

$$RMS = \sqrt{\frac{\sum (x_i - y_i)^2}{n}}$$

Hermes aplica estas distancias solamente al parámetro de frecuencia fundamental (F_0) e introduce una ponderación en función de la fuerza de la sonoridad vocal (calculada como la energía de la señal en el punto de F_0 considerado) con la finalidad de reforzar la importancia de las partes de la señal de palabra más susceptibles de transmitir una información importante. Rilliard/Aubergé (2003) han aplicado estas distancias a la comparación de medidas subjetivas y objetivas para una tarea de segmentación sintáctica de palabra, tomando en cuenta los tres parámetros de F_0 , duración silábica e intensidad, y sin utilizar ponderación. Los resultados obtenidos son conformes a las descripciones de la prosodia del francés de la literatura, con un predominio de las informaciones de F_0 y una importancia de las informaciones de duración, en tanto que la intensidad transmite poca información por lo que se refiere a la demarcación en unidades sintácticas.

Notamos enseguida que lo que denominamos aquí *distancias* consiste, de hecho, en (1) una medida de la cercanía de las formas de los contornos en el caso de la correlación y (2) una medida del alejamiento de las dos curvas en el caso de la distancia de los promedios cuadrados. De este modo, las dos curvas serán tanto más similares cuanto más cercana de 1 sea la correlación y de 0 la distancia de los promedios cuadrados.

medio de la Tabla I para un mismo hablante⁹): en efecto las dos variedades se diferencian poco en esta modalidad. El valor de 66,0% de intercorrelación obtenido para *I* es aún más bajo: las dos variedades se apartan mayormente en esta modalidad debido a un distinto *CTM* (cf. Romano 2001a, 2001b).

Tabla I.
Tablas de comparación entre los valores de intercorrelación obtenidos a partir de 5 repeticiones de dos frases correspondientes en dos códigos (variedad local = dialecto salentino) para un mismo hablante

Variété locale																	
Durée			Energie			F0											
	1	2	3	4	5		1	2	3	4	5						
1	1	0,838	0,731	0,972	0,836	1	1	0,938	0,980	0,979	0,936						
2	-	1	0,827	0,790	0,890	2	-	1	0,941	0,955	0,996						
3	-	-	1	0,673	0,623	3	-	-	1	0,964	0,924						
4	-	-	-	1	0,874	4	-	-	-	1	0,950						
5	-	-	-	-	1	5	-	-	-	-	1						
Intercorrélation moyenne						Intercorrélation moyenne						Intercorrélation moyenne					
0,805						0,956						0,924					
Italien régional																	
Durée			Energie			F0											
	1	2	3	4	5		1	2	3	4	5						
1	1	0,886	0,963	0,960	0,923	1	1	0,958	0,992	0,993	0,986						
2	-	1	0,882	0,858	0,806	2	-	1	0,956	0,950	0,946						
3	-	-	1	0,928	0,943	3	-	-	1	0,993	0,989						
4	-	-	-	1	0,909	4	-	-	-	1	0,991						
5	-	-	-	-	1	5	-	-	-	-	1						
Intercorrélation moyenne						Intercorrélation moyenne						Intercorrélation moyenne					
0,906						0,975						0,922					

Sobre la base de estas consideraciones, en Romano/Interlandi (2005) ya realizamos nuestras evaluaciones sobre la definición preliminar de *umbrales de intercorrelación* que, de manera más o menos dependiente de los estilos, se han manifestado para la duración en torno al 70%, para la energía en torno al 80% y para F₀ en torno al 90%.

⁹ Nótese que las repeticiones del otro hablante para la misma frase tienen una intercorrelación media del 92,2% (coherencia).

Tabla II.

Tablas de comparación entre los valores de correlación cruzada entre las 5 repeticiones de dos frases declarativas correspondientes en dos dialectos salentinos. La coherencia en este caso resulta del promedio de las medias

	2pcxtad1.txt	2pcxtad2.txt	2pcxtad3.txt	2pcxtad4.txt	2pcxtad5.txt	medie
1pcxtad1.txt	82,5%	85,4%	88,4%	72,7%	67,2%	79,2%
1pcxtad2.txt	77,9%	85,4%	85,9%	82,2%	75,8%	81,4%
1pcxtad3.txt	87,0%	80,6%	88,4%	83,3%	79,3%	83,7%
1pcxtad4.txt	68,5%	79,8%	75,8%	72,7%	64,2%	72,2%
1pcxtad5.txt	71,1%	80,2%	79,6%	75,6%	67,2%	74,7%
medie	77,4%	82,3%	83,6%	77,3%	70,7%	78,3%

Tabla III.

Tablas de comparación entre los valores de correlación cruzada entre las 5 repeticiones de dos frases interrogativas correspondientes en dos dialectos salentinos. La coherencia en este caso resulta del promedio de las medias

	2pcxtid1.txt	2pcxtid2.txt	2pcxtid3.txt	2pcxtid4.txt	2pcxtid5.txt	medie
1pcxtid1.txt	78,3%	78,3%	70,6%	63,3%	69,7%	72,0%
1pcxtid2.txt	63,8%	78,3%	61,4%	64,6%	57,9%	65,2%
1pcxtid3.txt	52,6%	64,9%	70,6%	60,9%	55,7%	61,0%
1pcxtid4.txt	62,4%	71,1%	64,0%	63,3%	60,2%	64,2%
1pcxtid5.txt	67,4%	75,7%	61,0%	65,2%	69,7%	67,8%
medie	64,9%	73,7%	65,5%	63,5%	62,6%	66,0%

2.2. Distancias prosódicas: aplicación a distintas variedades

La aplicación del cálculo de estas distancias a dos variedades que presentaban perfiles semejantes pero con una alineación temporal diferente nos ha permitido evaluar en primera aproximación la sensibilidad del índice a este tipo de diferencias (Miotti/Romano 2008). El estudio de las variedades consideradas en el presente trabajo nos ha permitido realizar unas evaluaciones más sobre la sensibilidad del índice introducido por otras fuentes de variación, como por ejemplo –para una misma modalidad en una misma variedad, en los datos producidos por un mismo hablante– por la diferente posición de un acento.

En los datos del hablante de la primera variedad considerada (Codroipo, Friuli, 820) hemos calculado, por ejemplo, la distancia entre las medias de tres repeticiones para las seis frases codificadas en la Tabla IV, caracterizadas por SN2 derivantes de la combinación de diferentes tipos acentuales:

(kw)fa: fdd+fdd - *proparossitono+proparossitono*,
 (kw)ja: ddf+dfd - *ossitono+parossitono*,
 (kw)sa: dfd+fdd - *parossitono+proparossitono*,
 (kw)va: fdd+dfd - *proparossitono+parossitono*,
 (kw)xa: ddf+fdd - *ossitono+proparossitono*,
 (kw)za: dfd+dfd - *parossitono+parossitono*.

Tabla IV.

Tabla de los valores de correlación (en %) entre las medias de tres repeticiones de seis frases declarativas con diferentes posiciones acentuales en el CTM (fa: fdd fdd; ja: ddf dfd; sa: dfd fdd; va: fdd dfd; xa: ddf fdd; za: dfd dfd)

	<i>kwfa</i>	<i>kwja</i>	<i>kwsa</i>	<i>kwva</i>	<i>kwxa</i>	<i>kwza</i>
<i>kwfa</i>	100	75	87	79	82	72
<i>kwja</i>	-	100	84	82	78	91
<i>kwsa</i>	-	-	100	68	89	88
<i>kwva</i>	-	-	-	100	61	75
<i>kwxa</i>	-	-	-	-	100	82
<i>kwza</i>	-	-	-	-	-	100

Los valores en negrita son los relativos a una diferencia de una sola posición localizada fuera del CTM y son, en efecto, entre los más elevados (> 85%), compatiblemente con la menor relevancia de estas diferencias (también en el nivel perceptivo). En otros casos, sin embargo, el valor de estas comparaciones –aun siendo todavía bastante elevado (> 80%)– puede bajar por debajo del que se obtiene comparando diferencias de dos posiciones (en cursiva). Al contrario, en un caso (*kwza* vs. *kwsa*), aun comparando dos enunciados con una diferencia en el CTM (aunque de una sola posición), el valor de correlación ha resultado de todos modos alto (88%). En los demás casos, las medidas de distancia han resultado inferiores al 80% (o bien sólo un poco superiores).

Para la misma frase analizada en los párrafos 2.1-2.3, en la modalidad I, podemos ahora pasar revista de las comparaciones cruzadas entre las frases de los hablantes varones de las tres variedades friulana de Codroipo (A), malagueña (B) y moldava de Iași (C), que aparecen ilustradas en la Fig. 1¹⁰.

En efecto, como han ilustrado Rilliard/Lai (2008), son las diferencias en los perfiles de las curvas (andamios parecidos en las secuencias de valores), más que las diferencias en valores absolutos entre las curvas, las verdaderas responsables de la baja de un índice basado en la correlación¹¹. Y es mediante esta medida como hemos evaluado las distancias entre las secuencias de valores presentes en los corpus AMPER

¹⁰ Valores negativos de correlación son obviamente índice de una mayor semejanza entre las secuencias de valores.

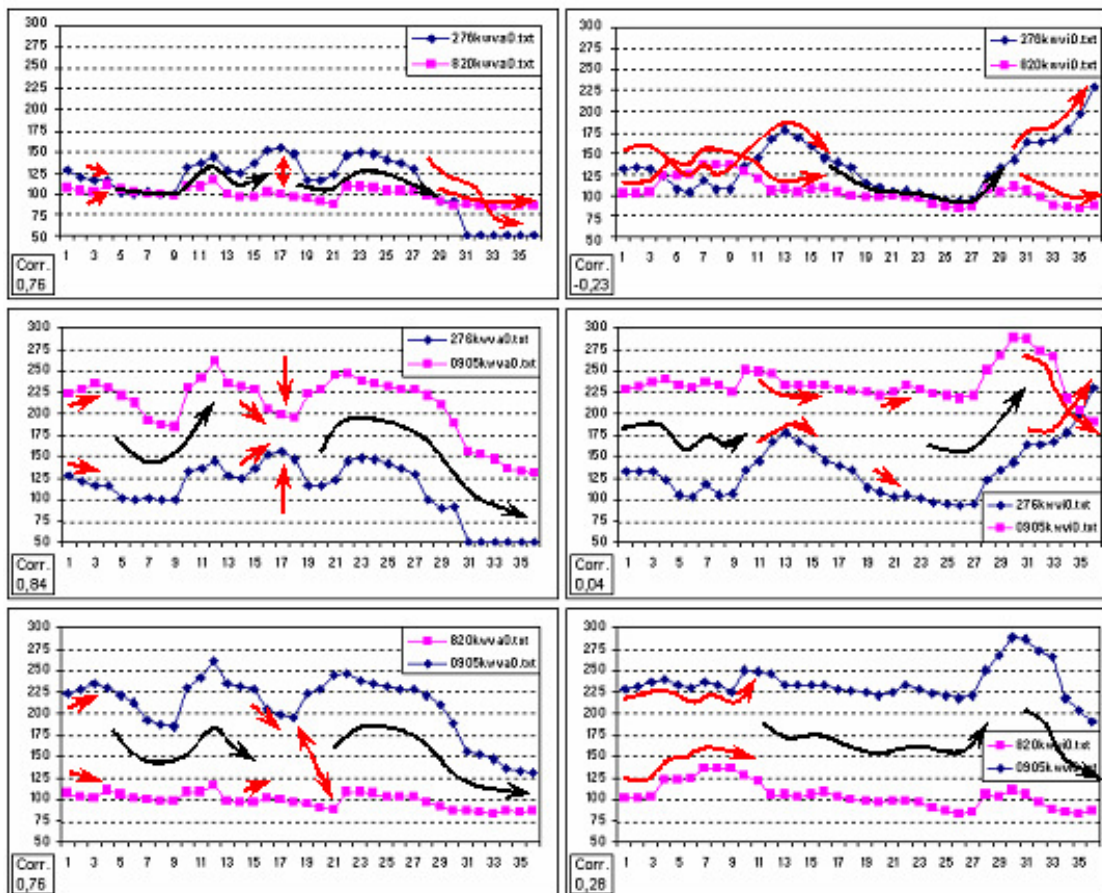
¹¹ Curvas parecidas, pero colocadas en registros diferentes, se medirían mejor por medio de una distancia basada en una diferencia cuadrática media. En nuestro caso, por razones obvias, esta medida es menos interesante (véase §1).

de las variedades que tomamos en consideración en este trabajo. En especial, calculando una media para cada tipo de frase (por ahora, excepto los ejemplos ilustrados arriba, se han considerado sólo aquellas con expansión a la izquierda), se han obtenido los intervalos de valores reproducidos en la Tabla V.

Asombrará constatar cómo las variedades, aun resultando bien diferenciadas entre ellas por este índice, cuyos valores son por lo general decididamente bajos, son distantes en base a medidas muy dispersas pero caracterizadas todas ellas por intervalos de variación más bien parecidos. Asombrarán también los valores negativos en la comparación entre los datos de A y C (hasta un mínimo de -11%), cuyas curvas no son, a fin de cuentas, completamente diferentes, y viceversa la mayor semejanza que se da en algunas secuencias de valores obtenidas para A y B (hasta un máximo de $+59\%$).

Fig. 1.

Gráficos de las comparaciones cruzadas entre las secuencias de valores medios de F_0 para las frases interrogativas de las tres variedades A (820), B (0905) y C (276) (con la indicación de la medida de correlación abajo a la izquierda de cada gráfico). Las flechas negras señalan las porciones de las secuencias de valores que se desarrollan siguiendo movimientos parecidos, en tanto que las flechas negras indican las porciones en las que se presentan particulares divergencias entre las curvas, muy probablemente responsables de la baja del valor de correlación.



Sin embargo, separando los valores de correlación entre frases *D* e *I* (véase para esta última modalidad los datos en la Tabla VI), observamos cómo los valores más bajos y los negativos se concentran mayormente en las medidas de intercorrelación entre los datos de A y B y de B y C (cf. ejemplos en la Fig. 1), mientras que resultan ligeramente más altos en las comparaciones A-C. Asombra, sin embargo, que, una vez más, en este caso el intervalo global de variación sea comparable con aquel de B-C (con un máximo que en ambos casos alcanza, de todos modos, un bajo 14%) y ello a pesar de que en los datos de estas variedades se registran secuencias de valores bien diferentes en su coda. La hipótesis es, por tanto, que estas medidas quedan invalidadas por una evaluación global de correlación que no tiene en cuenta las posiciones perceptivamente salientes para una evaluación impresionista de las semejanzas y de las diferencias entre las variedades. Una solución actualmente en discusión (también subrayada por Lai/Rilliard 2008) está relacionada con la necesidad de una ponderación para las partes perceptivamente relevantes.

Tabla V.

Tabla de las distancias prosódicas (en %) entre las variedades A, B y C (frases D e I).

	A	B	C
A	100	-5÷+59	-11÷+47
B	-	100	3÷42
C	-	-	100

Tabla VI.

Tabla de las distancias prosódicas (en %) entre las variedades A, B y C (sólo frases I).

	A	B	C
A	100	-5÷-11	4÷14
B	-	100	3÷14
C	-	-	100

3. Conclusión y perspectivas

En este trabajo hemos tratado de ilustrar un posible empleo de una medida cuantitativa en la evaluación de la distancia prosódica entre dos o más variedades.

Las medidas de distancia que proponemos, si no reemplazan el análisis experto del dialectólogo, permiten medir algunas divergencias de manera fiable y proporcionan unas indicaciones sobre la distancia perceptiva que podemos esperarnos entre dos dialectos. En nuestro caso, después de haber discutido de las posibilidades y de las modalidades de aplicación de la distancia propuestas en Romano/Miotti (2008) y en Rilliard/Lai (2008), hemos evaluado los resultados de una primera aplicación a los datos del corpus *AMPER*.

En términos generales, las frases consideradas difieren de forma considerable por sus perfiles melódicos en términos de amplitud y alineación de los movimientos.

Asimismo, la medida de intercorrelación propuesta, aun dando cuenta, de modo aceptable, de algunas propiedades globales, fracasa, sin embargo, en el intento de dar una medida cierta de la distancia prosódica percibida en ciertas porciones de enunciados particularmente significativas.

Por el contrario, algunos valores negativos en modalidad *I* se producen también en casos donde perceptiva y visualmente los perfiles no se alejan mucho, al menos en correspondencia con el *CTM*. Ello anima una vez más a redefinir estas medidas en términos multiparamétricos y perceptivos, con la posibilidad de ponderar partes especialmente relevantes a nivel lingüístico.

Es deseable que estos resultados sean confirmados por análisis realizados sobre la base de un número más importante de hablantes. En especial, resultaría muy útil calcular una distancia entre distintos hablantes del mismo punto de encuesta, con la finalidad de tener una idea más precisa de la especificidad de las estrategias propias de cada hablante.

La investigación sobre las distancias objetivas todavía está en sus comienzos y sería especialmente interesante empezar a aplicar este método de evaluación en mayor escala a la comparación de las distintas encuestas contenidas en la base de datos del proyecto *AMPER*. Ello debería permitir perfeccionar el modo de calcular estas distancias y de comprender mejor su eficacia. Esperamos que las posibles discusiones en el marco de este proyecto puedan brindar la ocasión de compartir diversas experiencias en vistas de la puesta a punto de un método cada vez más eficaz.

Bibliografía

- Contini, Michel (1992): "Vers une géoprosodie", en : *Actes du Nazioarteko Dialektologia Biltzarra Agiriak*. Bilbao: Real Academia de la Lengua Vasca, 83-109.
- ___ (2008): "Le projet AMPER: Passé, Présent et Avenir", en: Moutinho, Lurdes/Coimbra, Rosa Lúdia (eds.): *Actas das I Jornadas Científicas AMPER-POR*. Aveiro: Universidade de Aveiro, 9-19.
- Contini, Michel/Lai, Jean-Pierre/Romano, Antonio/Roullet, Stefania/Moutinho, Lurdes/Coimbra, Rosa Lúdia/Pereira Bendiha, Urbana/Secca Ruivo, Suzana (2002): "Un projet d'atlas multimédia prosodique de l'espace roman", en: Bernard, Bel/Marlién, Isabelle (eds.): *Proceedings of the International Conference Speech Prosody 2002*. Aix-en-Provence: Laboratoire Parole et Langage, 227-230.
- Moutinho, Lurdes/Coimbra, Rosa Lúdia (eds.) (2008): *Actas das I Jornadas Científicas AMPER-POR*. Aveiro: Universidade de Aveiro.
- Dorta, Josefa (ed.) (2007): *La prosodia en el ámbito lingüístico románico*. Santa Cruz de Tenerife: La Página Universidad.
- Gussenhoven, Carlos (2002): "Intonation and interpretation: phonetics and phonology", en: Bernard, Bel/Marlién, Isabelle (eds.): *Proceedings of the International Conference Speech Prosody 2002*. Aix-en-Provence: Laboratoire Parole et Langage, 47-57.
- Hermes, Dik J. (1998a): "Auditory and visual similarity of pitch contours", en: *Journal of Speech, Language and Hearing Research* 41, 63-72.
- ___ (1998b): "Measuring the perceptual similarity of pitch contours", en: *Journal of Speech, Language and Hearing Research* 41, 73-82.

- Lai, Jean-Pierre (2002): *L'intonation du parler de Nuoro (Sardaigne)*. Tesis doctoral en *Sciences du Langage*. Université Stendhal de Grenoble.
- Lai, Jean-Pierre/Rilliard, Albert (2008): "Outils pour le calcul et la comparaison prosodique dans le cadre du projet AMPER – l'exemple des variétés Occitane et Sarde", en: Turculeț, Adrian (ed.): *La variation diatopique de l'intonation dans le domaine roumain et roman*. Iași: Editura Universității Al. I. Cuza, 217-229.
- Lai, Jean-Pierre/Romano, Antonio/Rouillet, Stefania (1997): "Analisi dei sistemi prosodici di alcune varietà parlate in Italia: problemi metodologici e teorici", en: *Bollettino dell'Atlante Linguistico Italiano* (Torino) 21, 23-70.
- Lai, Jean-Pierre (ed.) (2005): *Projet AMPER - Atlas Multimédia Prosodique de l'Espace Roman, Géolinguistique*, Hors série 3.
- Martínez Celdrán, Eugenio (ed.) (2005): *Estudios de Fonética Experimental* 14. Barcelona: Laboratorio de Fonética. Universidad de Barcelona.
- Miotti, Renzo/Romano, Antonio (2008): "Una contribución a la comparación entre la entonación veneta y la andaluza", en: Pamies, Antonio/Amorós, M^a Cruz/Pazos, José Manuel (eds.): *Experimental Prosody, Language Design*, Special Issue 2, 91-98.
- Ohala, John (1983): "Cross-language use of pitch: an ethological view", en: *Phonetica* 40, 1-18.
- Ohala, John/Gilbert, John B. (1979): "Listeners ability to identify languages by their prosody", en: *Studia Phonetica* 18, 123-131.
- Pamies, Antonio/Amorós, M^a Cruz/Pazos, José Manuel (eds.) (2008): *Experimental Prosody. Language Design*, Special Issue 2. Granada: Granada Lingvistica.
- Rilliard, Albert/Aubergé, Véronique (2003): "Prosody evaluation as a diagnostic process: subjective vs. objective measurements", en: *International Journal of Speech Technology* 6, 409-418.
- Rilliard, Albert/Lai, Jean-Pierre (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: Moutinho, Lurdes/Coimbra, Rosa Lúcia (eds.): *Actas da I Jornadas Científicas AMPER-POR*. Aveiro: Universidade de Aveiro, 127-139.
- Romano, Antonio (1997): "Persistence of prosodic features between dialectal and standard Italian utterances in six sub-varieties of a region of Southern Italy (Salento): first assessments of the results of a recognition test and an instrumental analysis", en: Kokkinakis, George/Fakotakis, Nikos/Dermatas, Evangelos (eds.): *EuroSpeech 97. Proceedings of the 5th European Conference on Speech Communication and Technology*. Patras: University of Patras, 175-178.
- __(2001a): *Analyse des structures prosodiques des dialects et de l'italien régional parlés dans le Salento (Italie): approche linguistique et instrumentale*. Lille: Presses Univ. du Septentrion.
- __(2001b): "Variabilità degli schemi intonativi dialettali e persistenza di tratti prosodici nell'italiano regionale: considerazioni sulle varietà salentine", en: Zamboni, Alberto/Del Puente, Patrizia/Vigolo, Maria Teresa (eds.): *La dialettologia oggi fra tradizione e nuove metodologie*. Pisa: ETS, 73-91.
- __(2003): "Un projet d'Atlas multimédia prosodique de l'espace roman (AMPER)", en: Sánchez Miret, Fernando (ed.): *Actas del XXIII CILFR*, vol. I, Tübingen: Niemeyer, 279-294.
- __(2004): "Indices acoustiques suprasegmentaux dans la caractérisation des langues romanes: identification de variétés linguistiques et description des traits prototypiques», en: *Atti del convegno MIDL 2004 "Identificazione des langues et des variétés dialectales par les humains et par les machines"*. Paris: École Nationale Supérieure des Télécommunications, 91-92.
- Romano, Antonio /Contini, Michel (2001): "Un progetto di Atlante geoproso-dico multimediale delle varietà linguistiche romanze", en: *Actes des XI^{èmes} Journées du GFS*. Padova: Unipress, 121-126.
- Romano, Antonio/Interlandi, Grazia Maria (2005): "Variabilità geo-socio-prosodica: dati linguistici e statistici", en: *Géolinguistique* (Hors série) 3, 259-280.

- Romano, Antonio/Miotti, Renzo (2008): "Distancias prosódicas entre variedades románicas", en: Turculeț, Adrian (ed.): *La variation diatopique de l'intonation dans le domaine roumain et roman*. Iași: Editura Universității Al. I. Cuza, 231-249.
- Turculeț, Adrian/Botoșineanu, Luminița/Minuț Ana M./Romano Antonio (2004): "L'intonation du roumain au sein du projet AMPER", en: *Bollettino dell'Atlante Linguistico Italiano*, 27, Torino (2003), 269-274.
- Turculeț, Adrian (ed.) (2008): *La variation diatopique de l'intonation dans le domaine roumain et roman*. Iași: Editura Universității Al. I. Cuza.
- Vaissière, Jacqueline (1995): "Natural explanations for prosodic cross-languages similarities", en: *Proceedings of the 13th ICPhS* (Stockholm), 654-657.
- Vaissière, Jacqueline/Boula de Mareüil, Philippe (2004): "Divers aspects de l'identification d'une langue ou d'un accent: du segmental à la prosodie", en: *Actes du colloque MIDL 2004 "Identification des langues et des variétés dialectales par les humains et par les machines"*. Paris: École Nationale Supérieure des Télécommunications, 1-5.