NEW METHODS OF CLUSTERING THE PROSODIC VARIATION OF ITALO-ROMANCE DIALECTS

Valentina DE IACOVO

ABSTRACT • As suggested by several authors (among those, Contini 1991), prosodic variation may be analysed in a dialectometrical perspective, leading to the construction of geoprosodic maps. Language atlases are in facts conceived as empirical database allowing a detailed documentation of specific language features by the collection of a consistent number of locations. Later, data are treated and classified according to different methodologies in order 'to abstract and visualise a basic pattern from the immense amount of data found in the language atlases' (Nerbonne & Kretzschmar, 2013). This kind of approach can be adopted for the investigation of the prosodic variation, allowing the clusterisation of different dialectal varieties. For this study, eight samples of Italian dialectal varieties have been considered: the data processing consisted in measuring the prosodic distance (Hermes 1998, Romano 2001) existing among the varieties, expressed by means of a correlation matrix. Values have later been grouped according to several hierarchical clustering techniques such as heat maps, dendrograms and phylogenic trees, in order to illustrate how the evaluated samples can be grouped according to peculiar prosodic variations, fostering the prosodic analysis towards a geoprosodic perspective.

KEYWORDS • Geoprosodic Variation; Dialectometry; Cluster Analysis, Italian Dialectal Varieties.

1. Studies on prosodic variation

Several studies conducted in the last thirty years (Canepari 1985, Endo & Bertinetto 1996), showed, by means of different approaches, the existence of a great prosodic variability related to a specific dialectal area of Italy.[1] The constant growing amount of researches on the prosodic varieties spoken in certain areas (Grice 1995, Gili Fivela 2004, Marotta 2005, among others), linked to the need of organising digital data, led to the creation of new methods to visualise the collected data, such as open-source multimedia database (Prieto *et al*., 2010-14), leaving aside, however, the role played by the dialectal *substratum* in the identification of prosodic variation.

In this direction, several works (Contini 1984, Romano 1997, Interlandi 2003, Felloni 2011) followed a common path, trying to describe the prosodic system of various dialectal areas according to a unique protocol based on the comparison of controlled speech. Plus, within the same framework, it has also been possible to map a global perspective of the multiple dialectal domains (Fernández Planas *et al*. 2011, Moutinho *et al*. 2011, Fernández Rei *et al*. 2014) through the data grouping (by means of heat maps, clustering) highlighting the correlations in terms of prosodic distances among the selected varieties.

[1] For a more detailed description of the dialectal areas of Italy, see Pellegrini (1977).

The concept of intonation dialectometry took inspiration from the several works of Séguy (1973) and Goebl (1983), who measured the lexical variation across dialectal areas using statistical methods. The idea of applying this methodology to prosodic studies has been later developed by Hermes (1998) and saw a first application in Romano (2001) for the evaluation of the intercorrelation in a sequence of repetitions uttered by the same speaker. Several works followed this methodology (Rilliard & Lai 2008, Romano & Miotti 2008) until the first cartographic application of the prosodic distance (Moutinho *et al*., 2011).

2. Methodology

The analysed data set consisted of eight samples from six different dialectal areas: two from the Gallo-Italian area (Genoa and Trento), two from the Tuscan area (Prato and Pisa), two from the Central area (Rome, Frosinone), one from the Upper southern (Taranto) and one from the Southern (Pollina). For each variety, only one speaker has been recorded, for a total of three women and five men. Their age varied from 25 to 53 years old.

The corpus employed in the analysis consisted of a series of sentences with a SVO structure, uttered in both declarative and interrogative modality. The syntactic structure[2] is composed by a subject (SN1), a verb (SV) and an object (SN2); the SN1 is a proparoxytone trisyllabic word (preceded by a monosyllable, generally an article), while the SV is a paroxytone bisyllabic word. The SV2 is always a trisyllabic (also preceded by a monosyllable, generally an article) with different stress patterns (oxytone, paroxytone, proparoxytone)[3] in order to show how the stress structure affects the prosodic pattern, as shown in table 1:

SN1	SV	SN2
article+trisyll. proparoxytone	bisyll. paroxytone	article+trysillabic paroxytone
la papera *(the duck)*	*mangia* *(eats)*	*la patata* *(the potato)*

Table 1. Example of the syntactic structure of the sentence "*la papera mangia la patata*".

Through a series of different scripts, the utterances have been segmented and the main prosodic cues (f_0, duration, intensity) of each vocalic portion have been extracted. In a second moment, the f_0 values have been normalised and correlated among them applying a correlation measure. This is based exclusively on the fundamental frequency but weighted according to the signal energy associated to the point where the measurements are caught hence attaching importance to the voiced elements of a sentence, which are also considered as the most perceptively pertinent (Hermes, 1998) in terms of language identification. This measurement also ignores any idiosyncratic variations (such as register, main pitch) due to specific speaker qualities. An adaptation of the formula (Moutinho *et al*., 2011) readapted by d'Alessandro *et al*. (2011) from the works of Hermes (1998), is applied for the evaluation:

[2] Conventionally named *basic* structure.

[3] Sentences are named according to the specific stress typology of the last word of the utterance: *k* (oxytone), *t* (paroxytone) and *p* (proparoxytone).

$$r_{f_1 f_2} = \frac{\sum_i w(i)(f_1(i) - m_1)(f_2(i) - m_2)}{\sqrt{\sum_i w(i)(f_1(i) - m_1)^2 \sum_i w(i)(f_2(i) - m_2)^2}}$$

Figure 1. Formula calculating the prosodic distance taken by Moutinho *et al*. (2011).

f1 and *f2* represent the f0 values of the two intonation contours (expressed in semitons), while *m1* and *m2* the average values of these contours on the whole sentence. *W* is the weighting due to the signal energy, calculated as the average of the two energy values, measured at a given point for the two compared sentences (expressed in dB), while index *i* varies from 1 to the considered f0 points of the chosen sentence. This correlation measurement, obtained by specific scripts, allows the generation of a correlation matrix. The graphical result offered by the intonation curves is therefore associated to the equivalent matrices allowing a deeper insight on the cross-correlation of selected sentences by highlighting the degree of correlation (from a minimum of 0 –lower correlation- to a maximum of 1 –higher correlation-) among each sample. Values are therefore processed by a cluster analysis, specifically using several statistical outputs: the heatmap, for example, converts the varieties correlation in shades of red (max. correlation), blue (min. correlation) and white (low correlation) and group the potential clusters by a dendrogram (as shown in figure 2).

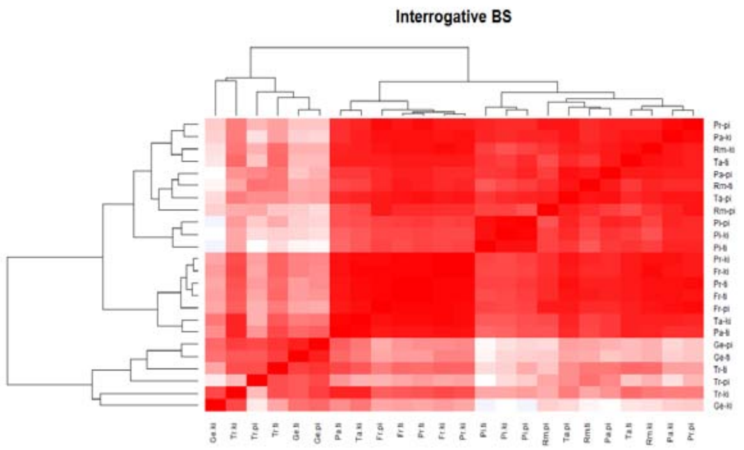

Figure 2. Cluster analysis of the interrogative modality for the eight samples: the first two letters identify the variety (*Ge*, *Tr*, *Pi*, *Pr*, *Rm*, *Fr*, *Ta*, *Pa*), while the last two specify the stress pattern (*ki*, *ti*, *pi*).

Figure 2 shows a neat distinction between the two Gallo-Italian varieties (Trento and Genoa) and the others, highlighted in the heatmap by pale shades and the dendrogram. In the second group, a further distinction is clear: the first one concerns the varieties of Frosinone and Prato (respectively divided according to the morphosyntactic structure, oxytone or paroxytone), the other one assembling two different varieties (Taranto and Pollina). In the second group, there is a first division including the Pisan variety and two main subgroups: the proparoxytone structures of Rome, Taranto and Pollina and the oxytone structures of Rome and Pollina.

3. Data analysis

The matrices allow to group the varieties in terms of prosodic distances, by means of a divisive dendrogram. Figure 3 shows for example the basic structures of all the varieties both in declarative and interrogative modalities. A first division concerns the Gallo-Italian varieties of Genoa (Ge) and Trento (Tr). While the Genoa declarative are well distinguished, the questions are grouped with the ones of Trento; interestingly, the oxytone (ka/ki) structures of both varieties are still separated from the paroxytone (ta/ti) and proparoxytone (pa/pi). In the other group, a first division concerns the variety of Pisa (Pi) and some declarative sentences of Pollina (Pa) and Rome (Rm), while in the other group, the more conspicuous; it is possible to distinguish the declarative sentences of Rome, Taranto (Ta), Prato (Pr), Frosinone (Fr) from their equivalent interrogative. In spite of the high number of data points in the chart, some subtle distinctions are still evident: some structures of south varieties (Ta-ki/Pa-ti, Ta-pa/Ta-ta) and Prato and Frosinone (Fr-ki/Pr-ki, Fr-ti/Pr-ti) are grouped together. Conversely, this is not the case for other examples (Pr-pi/Pa-ki).

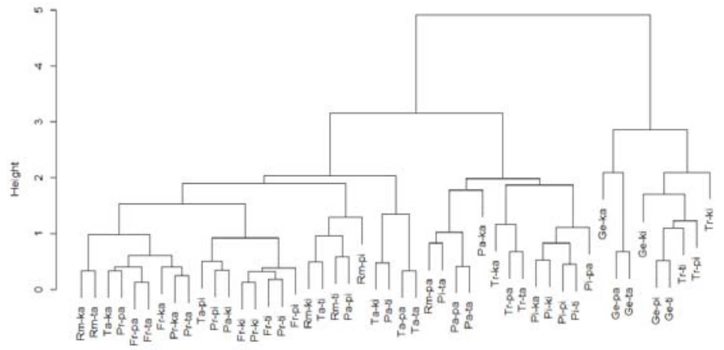

Figure 3. Dendrogram of the basic structures of all the varieties (declarative and interrogative modality).

Examining only the interrogative modality allows to focus on more specific details about the eight samples. Genoa and Trento are still considered closer compared to the others (always with a distinction between oxytone and paroxytone/proparoxytone); the Pisan variety, although belonging to the other subgroup, is, however, distinguished from the other varieties. Frosinone and Prato are grouped according to their stress pattern (oxytone and paroxytone), while the Romanesco variety seems more related to the southern ones. Another type of cluster analysis is the phylogenetic tree (see Baayen, 2008), leading to a graphic representation as the one shown in figure 4.

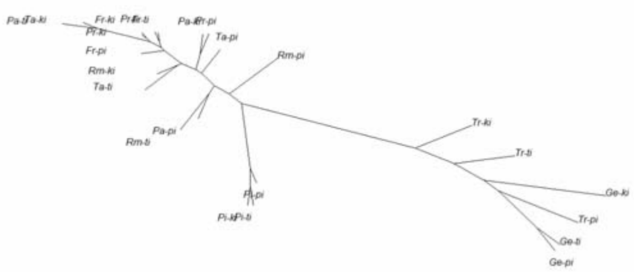

Figure 4. Phylogram of the basic structures of all the varieties (interrogative modality).

The phylogenetic tree in figure 4 shows a first distinction between the samples of Trento and Genova (on the right) and the other areas (on the left). The Pisan variety is quite detached from the other varieties, which appear more uniformed. Therefore, compared to the dendrogram representation, this kind of representation seems less useful in order to identify the prosodic correlation among the samples examinated.

4. The *DTW* analysis

The data have also been analysed by the dynamic time warping (*DTW*), a method employed for the classification of animal sounds and automatic speech recognition (Sakoe & Chiba 1978, Gamba *et al*. 2015). This kind of analysis considers the alignment between two-time series, exploiting temporal distortion between them. The sound similarity has been assessed by means of an algorithm based on the *DTW* and therefore a cluster analysis has been applied in order to identify the potential pattern groups. A first analysis consisted in the clusterisation of the fundamental frequency values for all the structures as shown in figure 5.

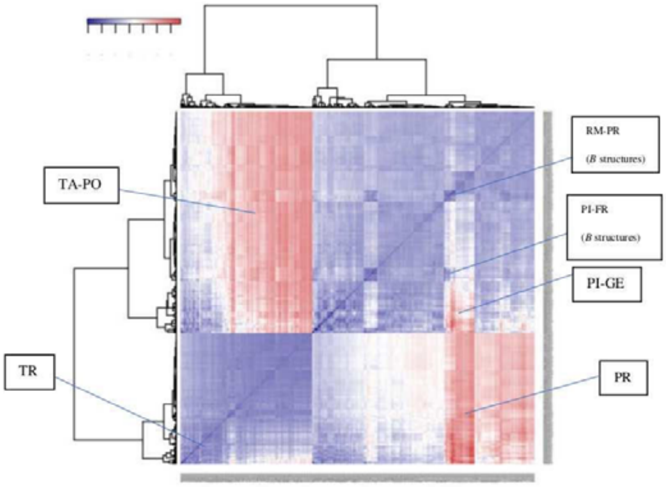

Figure 5. Heatmap generated by the *DTW* of the all the structures for all the varieties.

The heatmap in figure 5 shows some correlations among different varieties: for example, the Southern varieties (Taranto and Pollina) are grouped together, while the Trento variety, belonging to this main group, is however separated. On the contrary, the variety of Prato is well distinguished. This analysis also identifies some specific details such as the basic structures of Pisa and Frosinone and Rome and Prato.

5. Conclusions

The methodologies shown in these experiments represent several graphic representations based on different cluster analyses. The different dialectal varieties have been grouped according to specific prosodic features, which, through the correlation matrices, allow the representation of some samples of prosodic dialectal areas of Italy. Furthermore, results reveal an acceptable identification -and therefore, grouping- of the morphosyntactic structures, making this type of analysis sufficiently reliable. The dendrogram and the heath map represent the best output because they seem to highlight more clearly the distinctions among the analysed samples. Obviously, new dialectal samples are necessary to cover the remaining dialectal areas, in order to offer an overview of the geoprosodic variation occurring in Italy. This task should always consider the employment of reliable speech materials and the consideration of multiple prosodic features, aiming at describing also peculiar variations affecting the dialectal varieties.

Simultaneously, this will allow the comparison of Italo-Romance varieties in a broader perspective, also considering the whole Romance domain.

REFERENCES

Baayen, R.H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Canepari, L. (1985). *L'intonazione. Linguistica e paralinguistica*. Napoli: Liguori.

Contini, M. (1984). "L'intonation des phrases affirmatives et interrogatives, avec inversion en sarde: analyse et synthèse." In: *Bulletin de l'Institut de phonétique de Grenoble*, 13, 131-152.

Contini, M. (1991). "Où va la dialectologie?" In: *Quaderni di Semantica*, 2, 227-238.

d'Alessandro, C., Rilliard A., Le Beux S. (2011). "Chironomic stylization of intonation." In: *Journal of the Acoustical Society of America*, 129, 1594-1604.

De Dominicis, A. (2002). "Assertive e interrogative a Bologna e Roma." In: Regnicoli A. (eds.), *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*. Roma: Il Calamo, 129-136.

De Iacovo, V. & Romano, A. (2016). "La variation dialectale de l'intonation en Italie: le cas de Rome." In: *Dialectología, special issue 6 (Romance Geoprosody: Advances, Studies and Tools)*, 109-126.

Delattre, P. (1966). "Les Dix Intonations de base du français." In: *French Review*, 40 (91), 1-14.

Endo, R. & Bertinetto, P.M. (1996). "Aspetti dell'intonazione in alcune varietà dell'italiano." In: *Atti delle VII Giornate di Studio del Gruppo di Fonetica Sperimentale*, Naples (Italy), 27-49.

Felloni, M.C. (2011). *Prosodia sociofonetica. L'italiano parlato e percepito a Parma*. Milano: Franco Angeli.

Fernández Planas, A.M., Roseano, P., Martínez Celdrán E., Romera Barrios, L. (2011). "Aproximación al análisis dialectométrico de la entonación entre algunos puntos del dominio catalán." In: *Estudios de Fonética Experimental*, 20, 141-178.

Fernández Rei, E., Moutinho, L. & Coimbra, R.L. (2014). "As entoacións galega e portuguesa: a fronteira á luz da dialectometría e da percepción." In: Sousa, X., Negro, M. & Álvarez, R. (eds.), *Lingua e identidade na fronteira galegoportuguesa*. Santiago de Compostela: Consello da Cultura Galega, 115-141.

Gamba, M., Friard, O., Riondato, I., Righini, R., Colombo, C., Miaretsoa, L., Torti, V., Nadhurou, B. & Giacoma, C. (2015). "Comparative analysis of the vocal repertoire of Eulemur: a dynamic time warping approach." In: *International Journal of Primatology*, 36(5), 894-910.

Gili Fivela, B. (2004). *The phonetics and phonology of intonation: the case of Pisa Italian*. Tesi di Perfezionamento, Scuola Normale Superiore di Pisa.

Goebl, H. (1983). "Eléments d'analyse dialectométrique (avec application à l'AIS)." In: *Revue de Linguistique Romane*, 45, 349-420.

Grice, M. (1995). *The intonation of interrogation in Palermo Italian: implications for intonation theory*. Tubingen: Niemeyer.

Hermes, D.J. (1998). "Measuring the Perceptual Similarity of Pitch Contour's." In: *Journal of Speech, Language and Hearing Research*, 41, 73-82.

Hirst, D.J. & Di Cristo, A. (1998). *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press.

Interlandi, G.M. (2003). "La percezione dell'intonazione torinese: risultati di un test di riconoscimento." In: Marotta, G. & Nocchi N., (eds), *La coarticolazione. Atti delle XIII Giornate di Studio del Gruppo di Fonetica Sperimentale*. Pisa: Edizioni ETS, 193-201.

Marotta, G., (2003). "L'illusione prosodica." In: Atti del Convegno di Studi in memoria di T. Bolelli, *Studi e Saggi Linguistici*, XL-XLI, 237-258.

Marotta, G. (2005). "Toscane centrale et Toscane occidentale. Profils de l'intonation italienne." In: *Géolinguistique*, 3, 241-257.

Moutinho, L., Coimbra, R.L., Rilliard, A. & Romano A. (2011). "Mesure de la variation prosodique diatopique en portugais européen." In: *Estudios de Fonética Experimental*, 20, 33-55.

Nerbonne, J. & Kretzschmar, W.A. (2013). "Dialectometry++." In: *Literary and Linguistic Computing*, 28 (1), 2-12.

Pellegrini, G.B. (1977). *Carta dei dialetti d'Italia*. Pisa: Pacini.

Prieto, P., Borràs-Comes, J. & Roseano, P. (2010-2014). *Interactive Atlas of Romance Intonation*. Web page: http://prosodia.upf.edu/iari/.

Rilliard, A. & 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 parlers romans." In: Moutinho, L. & Coimbra, R.L. *Actas I Jornadas Cientificas AMPER-POR*. Universidade de Aveiro, Portugal, 127-139.

Romano, A. (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." In: *Proceedings of EUROSPEECH 1997*, Rhodes (Greece), 175-178.

Romano, A. (2001). *Analyse des structures prosodiques des dialects et de l'italien régional parlés dans le Salento (Italie): approche linguistique et instrumentale*. Lille: Presses Université du Septentrion.

Romano, A. & Miotti, R. (2008). "Distancias prosódicas entre variedades románicas." In: Turculet, A. (eds), *La variation diatopique de l'intonation dans le domaine roumain et roman*. Iaşi: EUAIC, 231-249.

Roseano, P. (2016). "Dos décadas de dialectometría entonativa." In: A. Iglesias, A. Romero, A. Ensunza (eds.), *Linguistic Variation in the Basque language and Education II*. Bilbao: Universidad del País Vasco, 56-80.

Sakoe, H. & Chiba, S. (1978). "Dynamic Programming Algorithm Optimization for Spoken Word Recognition." In: *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26, 1, 43-49.

Séguy, J. (1973) La dialectométrie dans l'Atlas linguistique de la Gascogne. In: *Revue de linguistique romane*, 37, 1-24.

VALENTINA DE IACOVO • PhD in Digital Humanities at the Department of Foreign Languages, Literatures and Modern Cultures, University of Turin. She works on the Italian section of the International AMPER (*Atlas Multimédia Prosodique de l'Espace Roman*) project and her main researches concern the geoprosodic study among dialects of Italy and the organisation of speech archives.

E-MAIL • valentina.deiacovo@unito.it