

Italianization processes in the Piedmontese dialects of Biella: a quantitative description based on *Multiple Correspondences Analysis*.

Within the Italian situation of *lingua cum dialectis*, one of the most intriguing cases of language contact is represented by the *italianization of the dialects* (ID). Under the pressure of the national language, Italo-Romance dialects tend to the reduction of their structural distance with Italian, under all the levels of analysis (see Sobrero 1997, Ricca 2010, Scivoletto 2014 and Cerruti 2016). To investigate how deep is the ID process in respect to the morphology and the syntax of the dialects means to measure the internal vitality of the dialectal structures themselves and to test to which extent they are capable of reacting to the innovations coming from the Italian language (for a discussion of the concepts of *external* and *internal vitality*, see Berruto 2011).

The aim of this work is to show how different phenomena of ID can be conceived as characteristic of specific groups of speakers.

After a brief discussion of the two concepts of *italianization* and *convergence*, the methodology employed to analyse the data, namely *Multiple Correspondences Analysis* (MCA), will be presented. This is a statistical tool which is useful for the identification of systematic relationships between variables displaying a certain degree of correlation¹. The goal of MCA is to capture the main underlying dimensions responsible for the greatest amount of variation in the data as possible, grouping different features together (similarly to other statistical techniques such as the *Principal Component Analysis*). The analysis results in a coordinate system which is specific for the dataset under examination. These kinds of analysis represent quite a novelty in dialectological studies, since they are usually employed for the description of variation in standard languages (see for Cerruti/Vietti 2022 for spoken Italian).

The dataset employed for MCA comes from a corpus collected by the author in three different areas of the Piedmontese district of Biella, namely the urban centres and the two valleys of Elvo and Sessera, which are located respectively at the eastern and western parts of the area. 80 informants – 40 males and 40 females, – took part in the study: half of them live in the urban zone, and half in the country (20 from each valley). In order to detect the main on-going tendencies from the point of view of linguistic change, the informants have been divided into two classes based on their age: one half of them was under 40, and the other half was over 70. In this way, data can be interpreted following the apparent-time hypothesis. 49 phonetic, morphological, and syntactic variables were extrapolated from the dataset.

A careful examination of the results of the application of MCA to the data results in an effective representation of the main on-going tendencies in these varieties. Not only it is possible to visualize the variability in the data – and thus to isolate different social and geographical varieties on the basis of the co-occurrence of traits – but also to observe the seeds of linguistic change.

Features displaying the same distributional behaviour can be used to describe the varieties of different groups of speakers. For example, a bulk of traits indicating a certain degree of language attrition – e.g. lack of gender agreement in adjectives of the second type like *fort*, “strong”, from lat. FORTIS, IS, see (1) – can be linked to a specific class of young informants, which show the typical characteristics of *semi-speakers* (see Dorian 1981 and Grinevald/Bert 2011).

- (1) Cole caramele a la menta a son nen tant fòrt
Those candies of the mint CLS.3SG are not much strong.MPL
“Those peppermint candies are not so strong”.

¹ For the purposes of this work, I employed the software *R*, specifically the function *MCA* from the package *FactoMineR*.

Some features can be linked to the age of the informants (e.g. concordance of the verb *to be* with a postposed plural subject in existential sentences, which results to be a case of pattern replication from Italian and is characteristic of the younger speakers and contrast with the traditional solution in (2)).

- (2) A j' è doe bote ant al frigo
CLS.3SG LOC be.PRS3SG two bottles in the fridge
“There are two bottles in the fridge”.

Indeed, age can be seen as a good explanation for the distribution of many features, but it does not coincide with either of the two axis of the graph. Along the vertical axis, we can visualize socio-geographical variation, while the horizontal dimension mostly represents linguistic change: the interplay of these two dimensions is responsible for a 20% of the variation in the data. Of course it is not a big percentage, but it increases consistently if we group some features together or we rule out of the count the traits closer to the origin of the axes, which result to be unmarked and do not contribute to explain variability.

One may question if MCA really can tell us something new about the data, or “simply” gives a – statistically relevant – confirmation to a qualitative analysis of the corpus. The answer is that something new actually emerges from the application of MCA: the unmarkedness of a great bulk of traits and the particular positioning of the variety of the Valsessera, which results to be more innovative (and hence nearer to the urban dialects) than the one of the Elvo Valley, which turns to be more conservative. So, the application of MCA allows us to better comprehend the role of different features and speakers in the community repertoire, both in terms of variation and of linguistic change.

References

- Berruto, G. (2011), *Considerazioni conclusive*, in B. Moretti et al. (eds.), «Vitalità di una lingua minoritaria. Aspetti e prospettive metodologiche», OLSI, Bellinzona: 289-302.
- Cerruti, M. (2016), *L'italianizzazione dei dialetti: una rassegna*, in «Quaderns d'Italià» 21: 63-74.
- Cerruti, M. / A. Vietti (2022), *Identifying language varieties. Coexisting standards in spoken Italian*, in Beaman, K.V. / G.R. Guy (eds.), «The coherence of linguistic communities: orderly heterogeneity and social meaning», Routledge, London: 261-280.
- Dorian, N. C. (1981), *Language death. The life cycle of a Scottish Gaelic dialect*, University of Pennsylvania Press, Philadelphia.
- Grinevald, C., M. Bert (2011), *Speakers and communities*, in P. K. Austin, J. Sallabank (eds.), «The Cambridge Handbook of Endangered Languages», Cambridge University Press, Cambridge: 45-65.
- Ricca, D. (2010), *Italianizzazione dei dialetti*, in R. Simone (ed.), «Enciclopedia dell'italiano», Istituto dell'Enciclopedia Italiana Treccani, Roma: 711-713.
- Scivoletto, G. (2014), *L'italianizzazione dei dialetti come caso di convergenza linguistica*, in RID XXXVIII: 75-101.
- Sobrero, A. A. (1997), *Italianization of the dialects*, in M. Maiden, M. Parry (eds.), «The dialects of Italy», Routledge, London: 412-418.