The contribution of dialectometry to the study of the dialects of Italy.  
A case study on Tuscan

Simonetta Montemagni+, Martijn Wieling*, John Nerbonne*

+Istituto di Linguistica Computazionale “Antonio Zampolli”, CNR, Italy
*University of Groningen, CLCG, the Netherlands

simonetta.montemagni@ilc.cnr.it, wieling@gmail.com, j.nerbonne@rug.nl

For forty years quantitative methods have been applied to the analysis of dialect variation: these methods have focused mostly on identifying the most important dialectal groups using an aggregate analysis of the linguistic data (Séguy 1973; Goebl 1984; Nerbonne et al. 1999). While viewing dialect differences at an aggregate level certainly gives a more comprehensive view than the analysis of subjectively selected features, the aggregate approach has never fully convinced linguists who found that it fails to identify the linguistic distinctions among the identified groups with the underlying linguistic features. In his seminal work on the dialects of Italy, Michele Loporcaro (2009) includes dialectometry among the methods and instruments of the dialectologist while criticizing that it measures structural distances among dialectal varieties without passing through a rationalization of the linguistic structure. This is a criticism shared with other linguists and dialectologists, among them Schneider (1988), but more often than not, dialectometry has been ignored. For example there is no mention in Chambers, Trudgill and Shilling-Estes’ Handbook (2002, 2008).1

As illustrated in Wieling and Nerbonne (2015), the past five years have seen a great number of improvements in the field of dialectometry. It is worth mentioning here the development of techniques for identifying the most important individual linguistic features underlying aggregate dialect variation, which can be seen as an answer to the criticism by Loporcaro and others. Other important advances concern techniques to simultaneously analyze the role of geographic and socio-demographic factors conditioning dialectal variation and to assess linguistic change in dialects.

We will illustrate the extent to which the recent advances of dialectometry can help to gain insight into the nature of linguistic variation – both synchronically and diachronically - in the study of the dialects of Tuscany, which have a special status in the complex puzzle of Italian dialects. This will be done by discussing the results achieved in a case study carried out over the last five years based on the corpus of dialectal data of the Atlante Lessicale Toscano (‘Lexical Atlas of Tuscany’, henceforth ALT, Giacomelli et al., 2000), a regional linguistic atlas focusing on dialectal variation throughout Tuscany, a region where both Tuscan and non-Tuscan dialects are spoken. ALT interviews were carried out in 224 localities of Tuscany, with 2193 informants selected with respect to a number of parameters ranging from age and socio-economic status to education and culture, on the basis of a questionnaire of 745 items. For the specific concerns of this study we used the web version of the atlas, ALT-Web, available at http://serverdbt.ilc.cnr.it/altweb/ (Cucurullo et al. 2006).

Starting from aggregate analyses of the whole ALT dataset, patterns of phonetic and lexical variation were reconstructed (Montemagni, 2007). If on the one hand the results achieved at the level of lexical variation turned out to be in line with the classifications of Tuscan dialects proposed in the literature by Giacomelli (1975) for what concerns the lexicon, on the other hand the patterns of phonetic variation that emerged did not appear to conform to the analyses by Giannelli (2000) and Pellegrini (1977). This result was unexpected and required further analysis aimed at gaining insight into the nature of phonetic and lexical variation in Tuscany and at investigating whether and to what extent the two resulting pictures correlated (Montemagni, 2008). To this end, we applied the technique of hierarchical spectral partitioning of bipartite graphs (Wieling and Nerbonne, 2010, 2011) to simultaneously identify synchronic patterns of dialectal variation and uncover the underlying linguistic features: this was done for both phonetic and lexical variation (Montemagni et al., 2012, Montemagni and Wieling, to appear). The same method of spectral partitioning of bipartite graphs applied to synchronic dialectal data was also used to investigate diachronic processes, thus contributing to a deeper understanding of the relationship between synchronic variation and diachronic change. By operationalizing Bärtol’s areal “norms” (Bartoli, 1925), the method was used to track the evolution of the spirantization phenomenon (called Tuscan gorgia) from different perspectives: geographically, phonologically and demographically (Montemagni et al. 2013).

1 The first edition of Chambers & Trudgill’s Dialectology (1980) was cautious about dialectometry (“utility […] not demonstrated”), but the second (1998) was positive and optimistic.
By revisiting the main steps of this case study carried out on Tuscan, we will focus on the contribution of dialectometry to the study of Italian dialects. The results achieved show that the most recent advances of dialectometry permit one to combine microscopic and macroscopic descriptions of dialectal variation within the same complex and articulated picture: new research questions can be answered, thus opening up new horizons of research. We will conclude by illustrating current directions of research which include the analysis of dialectal variation patterns across semantic domains and across generations.

References