Abstract
A careful investigation of synchronic patterns of linguistic variation with underlying linguistic features can lead to important insights into the comprehension of diachronic phonetic processes. In this paper, we showed that the method of spectral partitioning of bipartite graphs applied to synchronic dialectal data can effectively and reliably be used to investigate diachronic processes, thus contributing to a deeper understanding of the relationship between synchronic variation and diachronic change. This was illustrated through a case study carried out on Tuscan dialects, focusing on so-called Tuscan “gorgia”, a lenition process consisting of the spirantization of stop consonants. In particular, from a quantitative analysis of the sound correspondences involving voiceless and voiced stops, we tracked the evolution of the spirantization phenomenon in several respects. First, we tracked spirantization geographically, across Tuscany from the influential center of Florence to the peripheral areas. Second, we tracked it phonologically, from voiceless to voiced stops, and within each voicing class from velars to dentals and then to bilabials. Finally, we tracked it demographically, with young speakers using the most innovative sound correspondences more than old speakers. The fact that these results are in line with the literature on the topic of Tuscan “gorgia” demonstrates the potential of the method of spectral partitioning of bipartite graphs with respect to the reconstruction of diachronic processes starting from diatopically distributed synchronic dialectal data.
1. Introduction

The relationship between synchronic variation and diachronic change represents an issue debated since the early times of linguistics and dialectology which nowadays is attracting renewed attention as testified by recent studies using dialects as a testing ground for investigating language change (De Vogelaer and Seiler, 2012). An important line of inquiry in this direction was represented by the work of the Neolinguistic school, founded in the 1920s by the Italian linguist Matteo Bártoli (Bartoli, 1925; Bonfante, 1947), whose main outcome can be identified in the definition of “areal norms”, i.e. rules of thumb to be used for establishing the relative chronology of geographically competing linguistic variants. The three norms which were genuinely geographical are reported below, adapted from Trudgill (1975); given linguistic forms A and B:

1) if A is found in isolated areas, and B in areas more accessible for communication, then A is older than B;
2) if A is found in peripheral areas and B in central areas, then A is older than B;
3) if A is used over a larger area than B, then A is older than B.

Whereas the first norm is a widely accepted one, the second and the third norms remain linked to the name of Bártoli. As Chambers and Trudgill (1998) put it, the chief problem of Neolinguistics was that the school attempted to work with these principles as if they were laws, whereas they simply represent tendencies. When treated as laws, the different areal norms can lead to contradictory results and many exceptions can be found to them. On the other hand, by taking them as guidelines, Bártoli’s areal norms have considerable validity. However, in order to exploit details of diatopic variation as a tool for establishing the relative chronology of linguistic variants, a quantitative analysis is needed: i.e. when tendencies and preferences dominate over rigid principles, quantitative methods are the answer. This is possible due to the fact that one of the advantages of studying diachronic processes through synchronic evidence is represented by the easy accessibility to the latter: compared to diachronic data, dialect data provide a high resolution picture which can be usefully explored with quantitative analysis techniques.
The present paper is aimed at contributing to the ongoing debate on the relationship between variation and change: starting from a synchronic dialectometric analysis of linguistic variation, we try to reconstruct the underlying diachronic processes. To this specific end, we use a newly proposed dialectometric technique of co-clustering (called “bipartite spectral graph partitioning”) advanced by Wieling and Nerbonne (2010, 2011). This technique identifies groups of dialects on the basis of the aggregate analysis of a large corpus of dialectal data and simultaneously reconstructs the underlying linguistic basis. Through this technique it is possible to understand which factors underlie the identified patterns of variation, the role played by each of them and their interaction. This technique, which was successfully used to understand the linguistic basis of observed synchronic patterns of linguistic variation for different languages (Dutch and Tuscan Italian; Wieling and Nerbonne, 2010, 2011; Montemagni et al., 2012), is now exploited to reconstruct diachronic processes. The reliability and effectiveness of this method was tested in a case study carried out on Tuscan dialects, focusing in particular on Gorgia Toscana, a well-known lenition in which speakers say [ˈbaxo], [ˈbaho] or even [ˈbao] for /ˈbako/ ‘worm’, or [paθaʊa] for /paˈtata/ ‘potato’ in word-internal position, as well as [la 'xaza], [la 'haza] or [la 'aza] for /la ˈkaza/ ‘the house’ in sandhi (i.e. across the word boundary).

The case study is based on the dialectal data of the Atlante Lessicale Toscano (‘Lexical Atlas of Tuscany’, ALT, Giacomelli et al., 2000). Even though the published atlas documents lexical variation, the atlas material contains phonetic transcriptions on which we based the present study. This dialectal dataset has already been used in previous dialectometric investigations of patterns of phonetic variation in Tuscany (Montemagni, 2007, 2008; Montemagni et al., 2012). In particular, Montemagni et al. (2012) apply the hierarchical spectral partitioning (of bipartite graphs) to the ALT dialectal dataset: their results helped gain insight into the nature of phonetic variation in Tuscany, by simultaneously providing a classification of dialectal varieties and their underlying linguistic basis. They demonstrate that patterns of phonetic variation in Tuscan dialects appear to be
strongly connected with spirantization phenomena. In the present study, we aim at investigating spirantization phenomena across Tuscany with the final goal of exploring whether and how diatopic linguistic variation, which is detected in a synchronic dialectometric analysis, can be used to shed light on its diachronic evolution. On the methodological side, this study gave us the opportunity to assess the impact of a contextualised representation of sound correspondences in tracking down the evolution and diffusion of phonetic phenomena (in this case, spirantization).

2. The “Gorgia Toscana”

The phenomenon commonly known as Gorgia Toscana (literally, ’Tuscan throat’, henceforth referred to as Tuscan gorgia or simply Gorgia) is a phonetic process belonging to the general class of lenitions, i.e., consonantal weakening phenomena. It refers to a regular weakening process still active synchronically but whose earliest reference dates back to the beginning of the sixteenth century\(^1\) that occurs in Tuscan varieties of Italian and consists in the spirantization of stops (Giannelli and Savoia 1978, 1979-1980; Hajek 1996; Cravens 2000; Giannelli 2000; Marotta 2001, 2008).

Tuscan gorgia is thus the name used for the intervocalic weakening of the voiceless stop consonants /p/, /t/ and /k/ taking the form of spirantization to [ϕ], [θ], and [x/h] respectively, none of which occurs in the consonant inventory of Italian. The spirantization of /k/ and /t/ can even extend as far as deletion. Tuscan gorgia prototypically occurs intervocally, both within and across words in continuous speech: stops followed by a glide or a liquid consonant can also undergo the weakening process. Gorgia examples involving voiceless stops are reported in rows 1-3 of Table 1 (for examples in word-internal position) and in Table 2 (across word boundaries). However, Gorgia effects extend beyond voiceless stops.\(^2\) Giannelli and Savoia (1978), Marotta (2001, 2008) and Sorianello (2001) all note that the voiced stops /b/, /d/ and /g/ are also involved in the process of
weakening, surfacing as [β], [ð], and [ɤ/ɥ]: see examples in rows 4-6 of Tables 1 and 2. Note that even in this case the resulting surface realizations are not part of the Italian consonant inventory.

<table>
<thead>
<tr>
<th>Spelling</th>
<th>Standard Italian</th>
<th>Tuscan</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>pepe</td>
<td>*/pepe/</td>
<td>['pepe]</td>
<td>pepper</td>
</tr>
<tr>
<td>schiacciato</td>
<td>/skja'tʃaːto/</td>
<td>[skja'tʃaːto], [skja'tʃaːto]</td>
<td>flat</td>
</tr>
<tr>
<td>amico</td>
<td>/a'miko/</td>
<td>[a'miko], [a'miho], [a'mio]</td>
<td>friend</td>
</tr>
<tr>
<td>abete</td>
<td>/a'betɛ/</td>
<td>[a'betɛ]</td>
<td>fir</td>
</tr>
<tr>
<td>dado</td>
<td>/'daːdo/</td>
<td>['daːdo]</td>
<td>dice</td>
</tr>
<tr>
<td>lago</td>
<td>/'lavor/</td>
<td>['lavor]</td>
<td>lake</td>
</tr>
</tbody>
</table>

Table 1: Gorgia examples inside words

<table>
<thead>
<tr>
<th>Spelling</th>
<th>Standard Italian</th>
<th>Tuscan</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>i piedi</td>
<td>/i 'pj̥edi/</td>
<td>[i 'pj̥edi]</td>
<td>the feet</td>
</tr>
<tr>
<td>la torta</td>
<td>/la 'torta/</td>
<td>[la 'torta]</td>
<td>the cake</td>
</tr>
<tr>
<td>la casa</td>
<td>/la 'kazə/</td>
<td>[la 'kazə],[la 'hazə]</td>
<td>the house</td>
</tr>
<tr>
<td>e beve</td>
<td>/e 'beve/</td>
<td>[e 'beve]</td>
<td>(s)he drinks</td>
</tr>
<tr>
<td>e dorme</td>
<td>/e 'dorme/</td>
<td>[e 'dorme]</td>
<td>(s)he sleeps</td>
</tr>
<tr>
<td>la gamba</td>
<td>/la 'gamba/</td>
<td>[la 'gamba]</td>
<td>the leg</td>
</tr>
</tbody>
</table>

Table 2: Gorgia examples across word boundaries

According to the literature since the beginning of the 20th century, Tuscan gorgia is a process exhibiting extensive variation in frequency and extent. Rohlf's (1930) and Hall (1949) observe that the geolinguistic extension of spirantization varies across the involved consonants: /k/ is affected in a wider area than /t/, which is in turn subject to spirantization over a larger area than /p/. Giannelli and Savoia (1978) record decreasing levels of spirantization for /k/, /t/ and /p/ respectively in Florentine speakers. Bafile (1997) reports that the occurrence of less spirantized (or
non-spirantized) forms becomes more frequent, passing from the velar to the dental, and then to the labial. Sorianello (2001) finds that /k/ is the primary target of the *gorgia*, followed in frequency by /t/ and /p/. Concerning the spirantization of /b/, /d/ and /g/, it should be noted that it shows a much more restricted diffusion throughout the region diatopically. It is worth noting, however, that as reported by Giannelli and Savoia (1978) /b/ and /d/ in intervocalic position are spirantized to [β] and [δ] at a variable but frequent rate, whereas the use of intervocalic [γ] for /g/ is usual and stable (the latter is occasionally replaced by a glottal [ɦ]). For both voiceless and voiced stops, Tuscan *gorgia* thus appears to be sensitive to the place of articulation.

Fig. 1 Physical map of Tuscany, with indication of the ten provinces

Let us consider now the geographic distribution of the reported variation in frequency and extent of *Gorgia*: as reported in Giannelli (2000), Tuscany is variably affected by this phenomenon. In the central area, corresponding to the provinces of Florence, Siena, Prato and Pistoia, this process is very active: both voiceless and voiced stops undergo spirantization by default. Around this area,
the process is less advanced: while in the northwestern areas of Tuscany (corresponding to the provinces of Lucca, Pisa and Livorno) as well as in the South (Grosseto) Tuscan *gorgia* is restricted to /kl/, /tl/ and /dl/, in the eastern area (i.e. Arezzo province) *Gorgia* occurs only with /kl. *Gorgia* is absent in the Northwest where non-Tuscan dialects are spoken (see below). As pointed out by Marotta (2008), this distribution shows a pattern very commonly found in the diffusion of phonological changes: Tuscan *gorgia* spreads from the influential center of Florence, traditionally viewed as the epicenter, in all directions (see Fig. 1). From Florence, the *gorgia* spreads along the entire Arno valley, losing strength closer to the coast. It is also present to some extent in the Northwest and the Northeast, with the Apennines representing the northern border of the phenomenon, as well as in Siena and further south (but not in far southern Tuscany).

The observed asymmetry in presence and extent of synchronic spirantization of both voiceless and voiced stop consonants is also reported to hold diachronically. Izzo (1972) provides evidence that velars spirantized at least several generations before non-velars did. Note further that diachronically spirantization of /b/, /d/ and /g/ appeared later than voiceless stop spirantization.

From what we have considered thus far, we can conclude that we are faced with a double spirantization hierarchy, holding both diachronically and diatopically: on the one hand the hierarchy of spirantization is /kl/-/g/ > /tl/-/d/ > /p/-/b/ (i.e. *Gorgia* affects the velar to a greater degree than it does the dental which is in turn affected more than the bilabial), while on the other hand spirantization is more extensive with voiceless stops than with voiced ones. Tuscan *gorgia* is far from being an obligatory rule. Acoustic studies performed by Marotta (2001) and Sorianello (2001) show that stops do, in fact, surface among the allophonic variants. This fact combined with the asymmetric distribution of spirantization phenomena across Tuscan consonants shows that Tuscan *gorgia* presents itself as a gradient process exhibiting rich variation, both diachronically and synchronically; in the synchronic domain variation is found diatopically but also diastratically (Giannelli and Savoia, 1978, 1979-1980).
3. The Data Source

3.1 The “Atlante Lessicale Toscano”

The *Atlante Lessicale Toscano* (henceforth, ALT) is an especially designed linguistic atlas in which dialectal data have both a diatopic and diastratic characterization. The adjectives qualifying this linguistic atlas in its name are “lexical” and “Tuscan”. ALT is lexical in the sense that its main focus is on lexic-semantic variation, but this does not preclude it from containing valuable information with respect to phonetic variation. ALT is Tuscan in the sense that it is a regional atlas focusing on dialectal variation throughout Tuscany, a region where both Tuscan and non-Tuscan dialects are spoken. ALT data were collected between 1974 and 1986 from 2,193 speakers who were first selected with respect to a number of parameters including age, socio-economic status, education and culture and who were then each asked 745 questions.

In ALT, each dialectal item is assigned different levels of representation organized in layers of progressively decreasing detail going from phonetic transcription to different levels of orthographic representations (Cucurullo et al. 2006). For this study, we focused on phonetic transcription and the normalized (orthographic) representation level, where the latter was useful for its elimination of the effects of productive phonetic processes even while representing distinct morphological (both inflectional and derivational) variants. To illustrate this more concretely, phonetic variants originating from spirantization or voicing phenomena such as [skja'tʃ:αθa] or [skja'tʃ:ada] are both assigned the same normalised form, i.e. *schiacciata*, whereas words such as [skja'tʃ:ata] (singular) and [skja'tʃ:ate] (plural) as well as [skjatʃ:ə'tina] (diminutive) are assigned distinct normalized forms.

The alignment of the different representation levels was exploited to automatically extract all phonetic variants (henceforth, PV) of the same normalized form (henceforth, NF). Due to the features of the ALT normalized representation level, a study based on the analysis of PVs of the
same NF should document only genuine phonetic processes, without interference from any other linguistic description level (e.g. morphology).

3.2 The Experimental Dataset

For the specific concerns of this study, we used ALT dialectal data in a somewhat peculiar way: namely, we started from the attested phonetic variants which were elicited from speakers for lexico-semantic purposes (see above) without any a priori phonetically-driven selection. Szmrecsanyi (2012) argues that compared to linguistic atlas material dialectal corpora yield a more realistic linguistic signal, with two main advantages. First, they provide graded frequency information better matching the perceptual reality of linguistic input than discrete atlas classifications. Second, while the atlas signal is non-naturalistic, metalinguistic, and more competence-based in nature, text corpora provide more direct, performance-based access to language. In order to reduce some of the inherent problems of atlas data we used lexical answers to questionnaire items organised by type (corresponding to the normalised form) to infer the production frequency of specific phonetic features as a proxy for their salience in determining the observed patterns of phonetic variation. However, by analyzing questionnaire data in a corpus-like way one of the main advantages of atlas-based studies, i.e. the representative areal coverage of collected dialectal evidence, can no longer be taken for granted. This potential problem was overcome by enforcing a minimal areal coverage threshold for what concerns the selection of NFs (see below).

In particular, data selection from the ALT dialectal atlas was carried out by combining linguistic and geographical criteria. With respect to the former, only nouns and adjectives were selected, attested either as single words or as multi-word expressions. Phonetic variability represented the other linguistic criterion on the basis of which we performed data selection. We required that the number of phonetic variants associated with the same NF range between a minimum of 5 and a maximum of 34 (in ALT, this is the maximum number of PVs associated with the same NF). Geographical criteria were concerned with a) the network of the locations investigated which was restricted to the 213 (out of the 224) locations where Tuscan dialects are
spoken and b) the areal coverage of selected NFs, which we required to be greater than or equal to 100 (out of 213) locations. The resulting experimental dataset was composed of 444 NFs (covering the 4.64% of the whole set of diatopically varying NFs) including a total of 502,799 phonetic variant tokens.

Since we did not know in advance whether the selected sample of 444 NFs with associated phonetic variants was representative of the whole set of NFs having at least two PVs attested in at least two locations (used in Montemagni, 2008), we measured the correlation between phonetic distances in the overall dataset and in the sample selected, which turned out to be very high ($r = 0.994$). It follows that the sample selected can be reliably exploited to study phonetic variation across Tuscany.

Since the analysis method proposed (see Section 4.1) is based on the comparison between PVs collected through fieldwork and their phonetic realization in a reference variety, the experimental dataset underlying this study also includes the phonetic transcription of the selected NFs according to standard Italian pronunciation.

4. Methods

4.1 Obtaining Sound Correspondences

In order to obtain sound correspondences (henceforth, SCs) linking the dialectal allophone with its realization in our reference variety (standard Italian), we aligned the phonetic variants with their phonetically transcribed standard counterpart. We obtained the alignments using an adapted Levenshtein algorithm (Levenshtein 1965). The standard Levenshtein algorithm obtains the alignment between two strings by minimizing the number of deletions, insertions and substitutions needed to transform one string into the other. For example, the Levenshtein distance between the standard and a dialectal realization of *albicocca* ‘apricot’ is 2, since we need two operations (one substitution and one deletion) as shown below.
Instead of the standard Levenshtein algorithm which assigns the same cost to substitutions involving similar sounds (such as [u] and [o]) as opposed to different sounds (such as [a] and [i]), we used a version which employs automatically determined segment distances based on Pointwise Mutual Information (PMI; Church and Hanks, 1990). This method was introduced by Wieling et al. (2009) and found to yield superior alignments as well as acoustically sensible sound correspondences (Wieling et al., 2012). The PMI method assigns small distances to sound correspondences which occur relatively frequently (i.e. more frequently than would be expected on the basis of the relative frequencies of the two individual sounds). Greater distances are assigned to sound correspondences which occur relatively infrequently. Given the example above, the PMI-based Levenshtein algorithm yields more sensitive sound distances:

<p>| Standard Italian | a l b i k ɔ k: a |</p>
<table>
<thead>
<tr>
<th>Putignano</th>
<th>a r b i ɔ k: a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.032 0.049.....</td>
</tr>
</tbody>
</table>

As multiple speakers were interviewed in every location, we used the most frequent phonetic variant as representative of all attested PVs for every normalized form.

As we are interested in spirantization in this study, we focus on phonetic correspondences involving both identical and non-identical segments, also including insertions and deletions, with a stop on the reference (standard) side and with either a plosive or a spirantized (including absent)
realization on the allophonic (dialectal) side. Sound correspondences are represented as follows: on the reference side we report the broad phonemic transcription (placed between slashes), whereas on the dialectal side the narrow allophonic transcription is provided, placed between square brackets (e.g. /k/:[h]).

Since in the ALT dataset the same sound correspondence could originate from different phonetic processes, Montemagni et al. (2012) in their initial investigation of the ALT dataset enriched the representation of the sound correspondences with contextual information. In this study, we follow their approach and for each sound correspondence we identify the left and right (single segment) context. We use a rough context, as we only distinguish consonants (“C”), vowels (“V”), semi-vowels (“B”), gaps (marked as “-” and corresponding to an insertion or a deletion) and the word boundary (marked as “_” for the first and last segment of a word). For example, the SC /k/:[-] (i.e. the [-] indicates a gap) in the alignment above is recorded as V/k/V:V[-]V. This indicates that there are vowels to the left and right of both /k/ and [-]. This enriched representation of sound correspondences is crucial in the framework of the present study focusing on the propagation of a specific phonetic process (spirantization) across different types of context in Tuscan dialects.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pieve Santo Stefano</td>
<td>0.0000</td>
<td>0.0593</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Anghiari</td>
<td>0.0000</td>
<td>0.0683</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Antignano</td>
<td>0.0000</td>
<td>0.0132</td>
<td>0.0219</td>
<td>0.0219</td>
</tr>
<tr>
<td>Rosignano Marittimo</td>
<td>0.0181</td>
<td>0.0090</td>
<td>0.0226</td>
<td>0.0136</td>
</tr>
<tr>
<td>Cecina</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0528</td>
</tr>
</tbody>
</table>

Table 3: Excerpt of the matrix with locations x sound correspondences

After obtaining the sound correspondences, we count the frequency of each of these. We normalize these values by dividing by the number of words whose alignments include the specific sound correspondence, as not all words are attested in every variety. The normalized frequencies are
stored in a matrix (exemplified in Table 3), where rows represent locations and the columns distinct (contextualized) sound correspondences.

This matrix is then used as input for the technique (explained below) we employ to cluster locations together with their characteristic sound correspondences.

4.2 Clustering Sound Correspondences and Locations Simultaneously

We use hierarchical spectral partitioning of bipartite graphs (Wieling and Nerbonne, 2010) to simultaneously identify the geographical clusters as well as their characteristic phonetic features in the ALT dataset. A bipartite graph has two sets of vertices (i.e. points, representing locations and sound correspondences) and a set of edges connecting vertices from one set to the other (i.e. an edge connecting two vertices represents the occurrence of a sound correspondence in a location).

By calculating the singular value decomposition (SVD) of the input matrix and repeatedly applying the $k$-means (with $k$ equals 2) algorithm to these results, a hierarchical clustering is obtained in which sound correspondences are clustered together with locations. The mathematical details of the procedure are discussed in Wieling and Nerbonne (2010).

Following Montemagni et al. (2012), we scaled all columns of the input matrix (see above) to between zero and one, to ensure that every sound correspondence was equally important, regardless of its absolute frequency. After applying the hierarchical spectral partitioning method to this scaled input matrix, we obtain a hierarchical clustering where locations are clustered together with the sound correspondences.

4.3. Determining the most important sound correspondences for every cluster

The hierarchical spectral partitioning method yields many sound correspondences in every geographical cluster. To identify the most important ones, we use the method introduced by Wieling and Nerbonne (2011) which is based on calculating the DISTINCTIVENESS and REPRESENTATIVENESS of each SC.
**REPRESENTATIVENESS** (of a sound correspondence) simply measures how frequently the sound correspondence occurs in the cluster. For example, if there are ten locations in the cluster and the sum of the normalized frequencies equals five, the representativeness is 0.5 (5 divided by 10).

In contrast, the **DISTINCTIVENESS** of a sound correspondence measures how frequently the sound correspondence occurs within as opposed to outside of the cluster. It also takes the relative size of the cluster into account. For example, if the SC does not occur outside of the cluster, the distinctiveness is maximal (i.e. 1; the sound correspondence perfectly distinguishes the cluster from the others), irrespective of the size of the cluster. Alternatively, if a cluster contains half of the total number of locations and half (or less) of the total sum of the normalized frequencies, the distinctiveness is minimal (i.e. 0; the sound correspondence does not distinguish the cluster at all).

The values of both distinctiveness and representativeness range between zero and one (in the cases we are interested in). In line with Wieling and Nerbonne (2010) we average the representativeness and distinctiveness to obtain the **IMPORTANCE** value (ranging between 0 and 1) for every individual sound correspondence.

5. **Results**

In this section, we report the results of applying the hierarchical spectral partitioning method to the selected dataset. In order to assess the contribution of contextual information in reconstructing the spreading of phonetic phenomena, two different experiments have been carried out, respectively based on 16 context-free and on 84 contextualized sound correspondences extracted from the alignments involving the 444 selected NFs. The comparison of these results is expected to shed light on the impact and role, if any, of contextual information in reconstructing the diffusion of spirantization across Tuscany.

5.1 **Geographical Results**
In Fig. 2, the maps show the geographic visualization of the clustering of Tuscan varieties into five groups obtained by using contextualized SCs (left map) and context-free SCs (right map).

Fig. 2 Geographic visualization of the clusters obtained with contextualized SCs (left map) and context-free SCs (right map). Different shades of darkness indicate different clusters, which also correspond to different steps in the generalization of Tuscan gorgia. See text for further remarks.

It is clear that the general pattern is similar in the two maps: the phonetic areas identified are arranged in an onion-like shape built around a big central area covering the province of Florence and propagating in different directions, towards the south (in the province of Siena), east (in the province of Arezzo) and west (covering the provinces of Prato, Pistoia, Lucca up to most part of Pisa and Livorno). Around this central area, there is an external layer.

However, major differences can also be observed between the two maps. Let us first focus on the central area. In the left map, based on contextualized SCs, the clusters identified correspond to different areas organized as follows: a white core area centered around Florence and expanding in all directions, a light grey discontinuous area comprising locations mostly surrounding the white core, and two smaller darker grey regions again at the periphery with respect to the core. In the right
map, based on context-free SCs, this corresponds to the white area with the light grey one (sparse locations around the white core) acting sporadically as a kind of transition zone. The reverse appears to hold at the level of the outer layer which has approximately the same geographical coverage in both maps with the following differences: while in the left map it presents itself as a compact area (black), in the right map it is fragmented into two subareas marked in black and dark grey.

More generally, the map on the left, based on contextualized SCs, distinguishes sub-regions of Tuscany more finely than the map on the right, which ignores the context of sound correspondences. Making SCs sensitive to the phonological context in which they occur leads to a finer-grained depiction of the SC’s distribution. This was to be expected, but we could not know that the more sensitive analysis would be robust enough to support interesting aggregations.

5.2 Linguistic Results

We turn now to a discussion of the linguistic features underlying the clusters just described. In both maps the central area, with Florence as epicenter, is characterized by spirantization phenomena, whereas the features underlying the outer layer correspond to retention of (some) plosives. Besides this general common trend, the most salient features underlying the clusters represented in the two maps differ significantly.

Let us first focus on the left map based on contextualized SCs. Recall that the bipartite partition method groups sound correspondences together with locations. So each sound correspondence is linked to a group of locations. In order to identify the most characteristic sound correspondences for each group of locations, we calculate the importance of each sound correspondence in its cluster. Table 4 reports the ranked SCs with associated importance scores for the three main clusters identified. Let us start comparing their underlying features. In the first column it can be seen that underlying the white core cluster there are SCs corresponding to spirantization phenomena involving voiced stops (/b/, /d/ and /g/) as well as /k/ in word initial
position. For example, $V/g/V:V[\gamma]V$ (i.e. a $/g:/[\gamma]$ correspondence in an intervocalic context) has the highest importance score in the core cluster (0.319). The second layer cluster (light grey), in contrast, is characterized by SCs (reported in the second column of Table 4) corresponding to prototypical Tuscan *gorgia*, i.e. involving voiceless stops in intervocalic context, with extensions to other contexts (e.g. word initial position). With respect to the outer layer (third column), corresponding to the black cluster, we observe mainly SCs with plosives on the dialectal side: the feature list includes only two spirantized SCs involving the voiceless stop $/k/$ with a less spirant outcome $[x]$ which are however assigned quite a low rank in the list of ordered sound correspondences (i.e. they appear in the 25th and 27th position, respectively).

<table>
<thead>
<tr>
<th>Core cluster (white)</th>
<th>Second layer cluster (light grey)</th>
<th>Marginal cluster (black)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $V/g/V:V[\gamma]V$ (0.319)</td>
<td>1. $V/t/V:V[\theta]V$ (0.192)</td>
<td>...</td>
</tr>
<tr>
<td>2. $V/d/V:V[\delta]V$ (0.281)</td>
<td>2. $/<em>p/V:</em>[\phi]V$ (0.164)</td>
<td>25. $/<em>k/V:</em>[x]V$ (0.134)</td>
</tr>
<tr>
<td>3. $/<em>k/C:</em>[h]C$ (0.210)</td>
<td>3. $V/p/V:[\phi]V$ (0.152)</td>
<td>26. ...</td>
</tr>
<tr>
<td>4. $/<em>k/V:</em>[h]V$ (0.126)</td>
<td>4. $V/p/C:[\phi]C$ (0.144)</td>
<td>27. $V/k/V:V[x]V$ (0.117)</td>
</tr>
<tr>
<td>5. $V/b/C:[\beta]C$ (0.112)</td>
<td>5. $/<em>t/C:</em>[\theta]C$ (0.130)</td>
<td></td>
</tr>
<tr>
<td>6. $V/t/C:V[\theta]C$ (0.130)</td>
<td>6. $/<em>p/B:</em>[\phi]B$ (0.128)</td>
<td></td>
</tr>
<tr>
<td>7. $/<em>p/B:</em>[\phi]B$ (0.128)</td>
<td>8. $V/k/V:V[h]V$ (0.112)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Ranked contextualized spirantization-related SCs with associated importance scores (between parentheses) for the three main clusters identified.

Consider now the ranking of features within each cluster identified. Concerning the white core cluster, it is interesting to note the ranking of features involving voiced stops in fact reflects the spirantization hierarchy defined with respect to the point of articulation (i.e. with velars being associated with greater importance with respect to dentals and bilabials). As to the second layer
cluster, SCs with /t/ and /p/ on the reference side (i.e. standard Italian) are assigned a greater importance than those involving /k/.

Let us now turn to the clusters of the map on the right, based on context-free SCs. The features underlying the white central area correspond to spirantization phenomena involving both voiceless and voiced stops. They are reported in Table 5 (first column) with the associated importance scores:

<table>
<thead>
<tr>
<th>Core cluster (white)</th>
<th>Second layer cluster (light grey)</th>
<th>Marginal cluster (dark grey)</th>
<th>Marginal cluster (black)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. /t/:[h] (0.500)</td>
<td>1. /k/:[x] (0.197)</td>
<td>1. /t/:[t] (0.196)</td>
<td>1. /k/:[k] (0.178)</td>
</tr>
<tr>
<td>2. /d/:[ð] (0.484)</td>
<td>2. /p/:[p] (0.148)</td>
<td>2. /h/:[b] (0.096)</td>
<td></td>
</tr>
<tr>
<td>3. /t/:[θ] (0.449)</td>
<td>3. /h/:[d] (0.094)</td>
<td>3. /g/:[g] (0.091)</td>
<td></td>
</tr>
<tr>
<td>4. /p/:[φ] (0.421)</td>
<td></td>
<td>4. /h/:[d] (0.094)</td>
<td></td>
</tr>
<tr>
<td>5. /b/:[β] (0.421)</td>
<td></td>
<td>5. /g/:[g] (0.091)</td>
<td></td>
</tr>
<tr>
<td>6. /g/:[γ] (0.405)</td>
<td></td>
<td>6. /h/:[d] (0.094)</td>
<td></td>
</tr>
<tr>
<td>7. /k/:[h] (0.259)</td>
<td></td>
<td>7. /g/:[g] (0.091)</td>
<td></td>
</tr>
<tr>
<td>8. /t/:[a] (0.178)</td>
<td></td>
<td>8. /h/:[d] (0.094)</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Ranked context-free spirantization-related SCs with associated importance scores (between parentheses) for the three main clusters identified.

The light grey cluster (covering sparse locations around the white core) is characterized by just one SC involving /k/ with a less spirant outcome (see the second column). Proceeding towards the external area, the features underlying the black and dark grey external clusters correspond to retention of the plosive, i.e., resistance to gorgia (see columns 4 and 3, respectively). In particular, whereas the most external cluster is characterized by a single SC only, which is /k/:[k], the ranked SCs of the dark grey cluster involve voiced stops as well as the remaining voiceless ones. In contrast to the results obtained starting from contextualized SCs, in Table 5 it can be noticed that
the ordering of features within each cluster does not appear to be related in any way to the spirantization hierarchies: SCs involving voiced and voiceless stops are mixed together; the same appears to hold with respect to the point of articulation.

By comparing the linguistic results obtained with and without contextual information it is now possible to better assess its role in the reconstruction of the spreading of spirantization phenomena across Tuscany. With contextualized SCs (left map), the spirantization area (with Florence as epicenter) is articulated into different clusters, inversely reflecting the evolution of the phenomenon across the Tuscan consonantal phonology. The most important features distinguishing the core cluster (white) correspond to the most recent spirantization phenomena involving voiced stops and voiceless ones (only /k/) in non-canonical contexts, whereas the cluster immediately surrounding the core cluster is characterized by spirantization of voiceless stops (intervocally and elsewhere). Using context information, the marginal area presents itself as a unique cluster, mainly characterized by retention of plosive outcomes.

The result obtained without context information combines a compact spirantization area (with Florence as epicenter), involving both voiceless and voiced stops, with a marginal area characterized by retention of plosives subdivided into two clusters: the more internal cluster is characterized by the retention of plosives for all stops except /k/, which is associated with the more external cluster. It is worth pointing out here that /k/ is the first stop originally affected by Tuscan *gorgia*,

The results sketched above show that context information plays a central role: sound changes are recognized to be conditioned by phonetic context, as we saw in the case of Tuscan *gorgia*. Methodologically, we note that the approach used in this study has successfully detected the influence of context automatically. Contextualized SCs enable the detection of a linguistically well-founded and articulated picture, both at the level of regional coherence and the underlying linguistic features. In particular, using contextualized SCs we were able to “reconstruct” the spreading of
spirantization phenomena through Tuscan consonantal phonology, i) by originally involving the velar stop /k/, then /t/ and /p/ and subsequently the voiced stops /g/, /d/, and /b/, and ii) through different types of contexts, i.e. intervocally in medial word position but also as the realization of a sandhi application. In fact, going from the core to the outer layer, the important features associated with each cluster inversely follow the spirantization hierarchy. While the core area is characterized by the most advanced and recent spirantization phenomena, in the other layers spirantization is progressively restricted to voiceless plosives, with only /k/ being involved with quite a low salience in the external layer. By gradually moving away from the core, we first observe clusters characterised by SCs involving /p/, /t/, then by /k/ with progressively fewer spirant outcomes. Without contextual information, a more static picture emerges, with a single spirantization cluster.

5.3 Assessing the Reliability and Validity of the Results

In order to assess further the reliability of the conclusions drawn in Sections 5.1 and 5.2, consider now the geographic distribution of SCs instantiating spirantization phenomena involving voiceless and voiced stops and their spirantized counterpart intervocally. Note that as far as /k/, /t/ and /g/ are concerned we reported only one of the possible outcomes (namely [h], [θ] and [ɤ]), meaning that the spirantization area of these segments can in principle be wider if all spirant outcomes are taken into account. However, we are reminded that the aim here is to assess the validity of results achieved in this study. In the maps in Fig. 3, the frequency of occurrence of each SC class is represented in terms of increasing darkness: areas with a high frequency of SCs are colored light grey, and areas with only low frequency SC’s are colored more darkly. The maps in the two rows report the distribution of voiceless and voiced spirantization, respectively, as these occur intervocally. One can see that the core area around Florence is characterized by the spirantization of both voiceless and voiced stops, whereas the area surrounding it is characterized by voiceless spirantization only.
From a comparison of the maps in Fig. 3 it can be seen that the areal distribution of these features broadly reflects the spirantization hierarchies as reported in the literature (see Section 2), in particular:

1. Voiceless spirantization (maps in the top row) is more widespread than voiced spirantization (maps in the bottom row);

2. Concerning the point of articulation, it can be seen within each row that spirantization affects the velars to a greater extent than it does the dentals, which in turn are affected more than the bilabials (velar > dental > bilabial), i.e., /k/ > /t/ > /p/ and /g/ > /d/ > /b/.

Whereas the contrast between voiceless and voiced spirantization emerges clearly in the two series of maps, the differences are more subtle when the comparison is carried out with respect to the point of articulation: in the latter case, they are concerned with the areal distribution but also with the frequency of occurrence of the phenomenon (represented in terms of darkness).
This global analysis of the phonetic features underlying the dialectal clusters identified is also in line with the primary texts on the topic of Tuscan gorgia (see Section 2). Tuscan spirantization originally arose in Florence and spread rapidly in different respects: geographically, by propagating from Florence in all directions, especially southward and westward; and phonologically, by originally involving the velar stop /k/, then /p/ and /t/ and only subsequently the voiced stops /b/, /d/ and /g/. This is even more evident by comparing the maps in Fig. 2 with the map in Fig. 4 from Giannelli and Savoia (1979-1980), showing the weakening (indicated by spirantization and/or lenition/voicing) of intervocalic /k/, /t/ and /p/ in Tuscany. The white central area corresponds to the spirantization area, whereas the contiguous area on the west side is characterized by the co-occurrence of spirantization and lenition/voicing phenomena suggesting that we are faced with two different patterns of spirantization.

Fig. 4 Map showing the weakening of intervocalic /k/, /t/ and /p/ in Tuscany from Giannelli and Savoia (1979-1980)
5.4 Old vs. Young Speakers

Results described in the previous sections show that spirantization in Tuscany is still a native and vital feature, quite resistant to standardization. Giannelli and Savoia (1978, 1979-1980) report that the recent accelerated development and spread of Florentine spirantization throughout Tuscany is increasingly typical of younger generations. This issue can be further investigated on the basis of the ALT dataset, by exploiting one of its main features, viz. the fact that in every location multiple speakers were interviewed (between 4 and 29) so that each PV is not only anchored to a given location, but also to a specific speaker. To investigate whether and to what extent Tuscan spirantization also spread demographically across generations, we grouped the speakers into an old age group (born in 1930 or earlier – 1930 was the median year of birth) and a young age group (born after 1930). For every age group, we used the phonetic variant testified by the majority of the speakers in the respective group. In this case, we focused on contextualized SCs only.

It turns out that the general clustering pattern is the same across the two age groups with minor but significant differences observed at the level of underlying features. The same typology of features underlies the major clusters for both young and old speakers, albeit with different frequencies associated with different individual features, which is therefore reflected both in the ranking and the importance score assigned to them. In particular, younger speakers appear to use more innovative SCs more than older speakers do, i.e. in the core spirantization cluster, SCs involving voiced stops /g/, /d/ and /b/ (as opposed to voiceless ones), and in the external spirantization cluster, SCs involving /p/ and /t/ (as opposed to only /k/). Different causes may be hypothesized as to what underlies this result. First, according to Giannelli and Savoia (1978, 1979-1980) the spreading of spirantization in young generations is particularly evident in careless, colloquial, fast speech, while ALT data were elicited on the basis of a questionnaire focused on lexico-semantic variation, so that careless, informal and emotive pronunciations are rarely reported.
In fact, besides specific lexically-oriented questionnaire items, the ALT data do not systematically record phonetic differences among language registers. Second, it should also be kept in mind that the phonetic transcription in ALT is coarse-grained, and finer distinctions between different spirantization degrees are not accounted for.

6. Conclusion

In this paper, we showed that the method of spectral partitioning of bipartite graphs applied to synchronic dialectal data can effectively and reliably be used to investigate diachronic processes, thus contributing to a deeper understanding of the relationship between synchronic variation and diachronic change. This was illustrated through a case study carried out on Tuscan dialects, focusing on so-called Tuscan gorgia, a lenition process consisting of the spirantization of stop consonants.

It turned out that a careful analysis of the sound correspondences involving voiceless and voiced stops provides valuable information for the reconstruction of the diachronic process of spirantization in Tuscany. In particular, starting from the synchronic patterns of variation identified with respect to stop spirantization it was possible to reconstruct the evolution of the phenomenon over time. This was done by looking at each cluster identified as corresponding to a time slice in diachronic change: following Bàrtoli’s areal “norms” (i.e., rules of thumb), in particular his second and third (focusing on central vs. peripheral distributions and larger vs. smaller ones), features underlying the clusters corresponding to peripheral areas were taken to represent the older developments. Proceeding towards the core cluster, progressively more recent developments were found, with the core area centered around Florence being characterized by the most recent developments of the Tuscan gorgia, namely spirantization of voiced stops. From this analysis it was also possible to trace the development of spirantization phenomena in Tuscany: one development starts with voiceless stops and then is extended to voiced ones; another development starts with velars, then proceeds to dentals and bilabials; yet a third perspective is concerned with the context
of occurrence of the phenomenon, developing from intervocalic contexts in word internal position, to other word-internal positions and then to contexts involving sandhi.

The evolution of the spirantization phenomenon was thus tracked in several respects. First, it was tracked geographically, across Tuscany from the influential center of Florence to the peripheral areas. Second, it was tracked phonologically, from voiceless to voiced stops, and within each voicing class from velars to dentals and then to bilabials. Finally, it was tracked demographically, with young speakers using the most innovative developments more than old speakers. The fact that these results are in line with the literature on the topic of Tuscan gorgia demonstrates the potential of the method of spectral partitioning of bipartite graphs with respect to the reconstruction of diachronic processes starting from diatopically distributed synchronic dialectal data.

On the technical side, this study gave us the opportunity to test the impact and role of a contextualised representation of sound correspondences which may lead to a better founded analysis of synchronic patterns of dialectal variation from which the evolution of diachronic processes could be reconstructed.

Acknowledgements
The research reported in this paper was carried out in the framework of the Short Term Mobility program of international exchanges funded by the National Council of Research (CNR, Italy). We also thank the anonymous reviewers for their comments which have helped to improve this paper. Thanks are also due to Alberto Montemagni who helped us with the pictures in the paper.
Notes

1. For historical details regarding Tuscan *gorgia*, the interested reader is referred to Izzo (1972).
2. Whereas the term *Gorgia* traditionally referred only to voiceless stop spirantization (see for instance Giannelli and Savoia, 1978 and 1979-1980 or Hajek, 1996), more recently (see e.g. Marotta 2008) it has been used to also refer to voiced stop spirantization.
3. ALT, originally published in Giacomelli *et al.* (2000), is currently available as an online dialectal resource at http://serverdbt.ilc.cnr.it/ALTWEB. This is the version used for this study.
4. This is the case for dialects in the north, namely Lunigiana and small areas of the Apennines (so-called Romagna Toscana), which belong to the group of Gallo-Italian dialects.
5. The phonetic alphabet used in the ALT project was a geographically specialized version of the “Carta dei Dialetti Italiani” (CDI) transcription system. For this study, the whole ALT corpus of phonetically transcribed data was converted to the International Phonetic Alphabet (IPA).
6. As in ALT verbal answers represented by different inflected forms are not always explicitly marked, verbs were excluded from the experimental dataset to prevent potential noise deriving from verbal morphology.
7. Note that multi-word expressions selected were represented by “frozen” word combinations, showing no variability due to the insertion/deletion of constituents.
References


