DOI: 10.17469/O2103AISV000006

CLAUDIA ROBERTA COMBEI, OTTAVIA TORDINI

A pilot sociophonetic study on open-mid vowels uttered by young male and female speakers of the Pisan variety

Based on a small dataset of spontaneous speech produced by 6 speakers aged 18-20, this study aims to acoustically describe vowel quality in the Pisan variety of Italian with respect to male-female differences. The Pisan variety is characterized by the lowering of [ɛ] and [ɔ], which is said to be a sociophonetic cue of vernacular speech. The initial aim was to explore whether gender-related sociophonetic differences influence the production of [ɛ] and [ɔ] in young speakers, and which of the two groups (males or females) were more likely to adopt this vernacular feature. However, we did not find consistent evidence of [ɛ] and [ɔ] lowering in the analyzed corpus. Nevertheless, there was variation in the production of [ɛ], which differentiated the male Pisan speakers from their Florentine counterparts. For this reason, this study also tested an automatic classification system of [ɛ] vowels, which could predict with decent accuracy (77%) the variety to which a specific vowel token belonged[1].

Key words: sociophonetics, acoustics, Pisan, open-mid vowels.

1. *Introduction*

> Exploring how physiological and social factors interact in the speech of both men and women is important in understanding the linguistic principles governing language variability.

As Byrd (1994: 41) highlights, gender-related features are considered fundamental to a proper understanding of language variability. However, a large number of studies in the field of experimental phonetics have placed more emphasis on male speakers, since they represent the unmarked choice[2]. On the other hand, female speakers have received less attention in phonetic research, both in acoustics and speech perception, for various reasons.

For example, Fant's early work (1960), which showed the source-filter interaction and paved the way for subsequent studies, was mainly based on the examination

[1] Authors' note: although the contribution is the result of the joint research activity of both authors, the responsibility for drafting the article and carrying out specific analyses has been distributed as follows: § 1, § 2.1 OT; § 2.2 and 3 CRC and OT; orthographic and phonetic transcriptions, segmentations and feature extractions were performed by both authors (vowel [a]: OT; vowels [ɛ] and [ɔ]: CRC and OT); in §4, the plot representations and their interpretations were made by CRC, while OT explained their sociophonetic implication; § 4.1, 5, 5.1 are to be attributed to CRC, § 6 to OT and CRC, and bibliography to OT.

[2] For a more specific and detailed discussion about this topic, see Ferrero, Magno Caldognetto & Cosi (1996).

of male physiological characteristics. Another reason behind the choice to study males rather than females may lie in the difficulty involved in identifying women's higher fundamental frequencies. In fact, this often hinders an accurate analysis of the location of formant frequencies, making it more difficult to detect phonetic contrasts (Klatt, Klatt, 1990).

As far as Italian regional varieties are concerned, until now, only a few studies have focused on gender differentiation related to speech production (Filipponio, Cazzorla, 2016; Nodari, 2016). More specifically, phonetic features of regional varieties of the Tuscan Western areas have been thoroughly investigated, but most of these analyses have not regarded gender as a relevant factor of variation (Calamai, 2001; Nocchi, Calamai, 2009).

2. *The study*

2.1 A brief overview of lowering of open-mid vowels in Western Tuscany

From a phonetic-phonological point of view, the Western area of Tuscany is characterized by specific phenomena which contribute to separating West (especially Pisa and Leghorn) from the rest of the region. In fact, the lowering of stressed open-mid vowels [ɛ] to [æ] and [ɔ] to [ʌ] has long been noted in the literature as a diatopically marked feature in this area (Giannelli, 2000; Calamai, 2001, 2003, 2004, 2009; Marotta, Calamai & Sardelli, 2004; Calamai, Ricci, 2005; Nocchi, Calamai, 2009). It can thus be considered relevant for the identification of the above-mentioned regional varieties of Italian, as well as being a cue of vernacularity (Giannelli, 2000; Calamai, 2001, 2004; Marotta, Calamai & Sardelli, 2004; Nocchi, Calamai, 2009). Many studies contributing to the analysis of the open-mid vowels lowering do not take gender into account (Calamai, 2001; Nocchi, Calamai, 2009). Moreover, most of them investigate isolated words or target words within carrier sentences[3] (Calamai 2001, 2003), excluding spontaneous productions.

2.2 Aims

Based on these assumptions, we carried out an acoustic analysis using both un-normalized and normalized vowel formant values on a small *ad hoc* collection of spontaneous and semi-spontaneous speech produced by young Pisan speakers. In order to provide an original contribution to the interpretation of this phenomenon, we considered gender differences within spontaneous productions, excluding written material, scripted dialogues and carrier sentences (Calamai, 2001). On that account, this study aims to make explicit at the sociophonetic level the link between gender and vernacularity, regarding variation in both diastratic and diatopic contexts. This research also tests a classification system that

[3] One notable exception is Nocchi, Calamai (2009). In this study, the authors made a comparison of spontaneous productions in Pisa, Florence, Arezzo and Leghorn. However, they analyzed the first two formants and the duration of stressed and non-stressed vowels, as well as the position of the token within the word. Our study concentrates only on the young generation of Pisan speakers, including both male and female subjects.

exploits data mining tools, using vowel formant values as features, in order to assess the extent to which vowel discrimination can be automated.

3. *Methods and tools*

In terms of settings and procedures, this kind of research requires a thorough review of various methodological issues. Several questions have been raised: the suitability of speech elicitation techniques; the choice of methods employed for acoustic analysis methods; and the statistical tools for data visualization and interpretation. In addition, it is often problematic to describe the place of articulation of vowels, because they produce a smaller and less consistent narrowing of the vocal tract than consonants (Calamai, 2004).

The first step consisted of building a small and homogeneous *ad hoc* corpus[4]. We recruited 6 young people, aged 18-20 years, who were born and had lived in Pisa all their lives. The sample was evenly divided between male and female speakers. Before each recording, subjects were asked to fill in a sociocultural questionnaire to collect information on their linguistic background, level of education, their parents' level of education and occupation, and their daily use of the dialectal variety. At an early stage of this study, education level and social status were not considered impactful. The Pisan speakers showed homogeneous characteristics, since all of them belonged to the same social network and social class (upper-middle). We chose young subjects for two reasons: firstly, because this approach allowed us to target the lowering of open-mid vowels in a uniform group; secondly, and most importantly, we wanted to examine whether there was experimental evidence supporting that these vowels were indeed lowering in young Pisan speakers.

Concerning the data, the speech was either spontaneous or semi-spontaneous, collected by means of the map-task elicitation technique and unstructured interviews. The use of the map-task was intended to increase the occurrence of the linguistic phenomena we were interested in, and allowed us to perform analyses both at an acoustic and sociolinguistic level. As for the sampling, the speakers were recorded with Praat[5] (Boersma, Weenink, 2016), using a Samson METEOR MIC cardioid pickup microphone (condenser diaphragms: 25mm). The sampling parameters were the following: mono channel, 16-bit, 16,000 Hz, linearly encoded WAV.

Due to numerous creaky voice productions in 2 out of 3 Pisan male speakers, at times formant measurements appeared to be significantly distorted. The open-mid vowels produced by the Pisan speakers were subsequently compared with those of a control group of four Florentine speakers. We carried out a contrastive analysis because the lowering of open-mid vowels has not yet been attested in the Florentine Italian, since it is mainly a feature of the Western area of Tuscany (Calamai, 2001; Sorianello, 2002). The Florentine data was retrieved from the CLIPS[6] corpus (Corpora e Lessici dell'Italiano Parlato e

[4] Differently from Nocchi, Calamai (2009), and because it did not suit the needs of this study, we did not use the material stored in Archivio del Parlato Italiano (API), retrieved from http://www.parlaritaliano.it/index.php/it/dati/40-api-archivio-del-parlato-italiano.

[5] Retrieved from http://www.praat.org/.

[6] Retrieved from http://www.clips.unina.it/it/corpus.jsp/.

Scritto), and consisted of audio recordings of 4 young speakers (2 males and 2 females). It was segmented and further analyzed, following the same procedures as for the Pisan informants. This type of analysis, also employed by Calamai (2004), allowed us to obtain a good comparison between the two varieties.

4. *Acoustic analysis*

The present study solely concerns stressed vowels[7], so as to reduce the number of dependent variables and to focus on the sociophonetic features of interest. The vowel extraction and the acoustic data processing were performed with Praat and ELAN[8]. The audio material was transcribed and segmented into three different Praat tiers: sentences, words and target vowels. We obtained 820 tokens of open-mid vowels: 555 for [ɛ] and 265 for [ɔ], roughly 50 tokens per speaker for [ɛ] and 20 tokens per speaker for [ɔ]. In addition, we also segmented stressed tokens of [a] (320 in total, ca. 30 per speaker), both in the Pisan and Florentine productions. Following the literature on the phonetic phenomena in Western Tuscany, [a] could be characterized by a more posterior place of articulation (Giannelli, 2000; Calamai, 2004). Thus, analyzing the acoustic features of this low vowel would help us estimate the vowel space for each speaker. Furthermore, it could contribute to targeting the lowering phenomenon in a more specific and defined space[9]. Using two *ad hoc* Praat scripts, we extracted the vowel duration as well as the first three formant frequencies and the fundamental frequency (f0) from the vowel midpoint. However, duration and f0 were eventually excluded from the final analysis due to the persistent variation among the spontaneous productions, mostly due to large variability in speech rate (Calamai, 2015), even within the same sentence. After the automatic extraction, we noticed that for a consistent number of tokens (ca. 25%) it was not even possible to detect f0 values. Therefore, the following values were calculated: single-point measurements, means, and standard deviations for the first three formants (F1, F2, F3) of the open-mid vowels. These parameters allowed us to plot the vowels using the R[10] package "vowels" (Kendall, Thomas, 2009).

For the first series of graphs, we used the raw formant values expressed in Hz. The graphs representing all tokens of [ɛ] uttered by all 10 speakers clearly show that the values reported for female speakers, both Florentine (efifem1, efifem2) and Pisan (epifem1, epifem2, epifem3), tend to be dispersed (see Figure 1). On the other hand, the Pisan male speakers appear more homogeneous, since the values reported for this group are noticeably similar. As far as [ɔ] is concerned, the data seem to suggest that the position of [ɔ] in the vowel space is comparable for most speakers, both Pisan and Florentine (see Figure 2). With regards to [a], all Pisan

[7] Differently from Calamai (2001).

[8] Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/.

[9] In the present study, we did not compute the Euclidean distance between [a] and [ɛ] and [a] and [ɔ]. However, this aspect could be further investigated in a next study.

[10] Retrieved from https://www.r-project.org.

male speakers (apima1, apima2, apima3) have similar values, unlike the females. In fact, only two Pisan females (apifem2 and apifem3) have similar productions, while the values for the other female speaker (apifem1) resemble those reported for male speakers (see Figure 3).

Figure 1 - *All tokens of [ɛ] uttered by all 10 speakers*

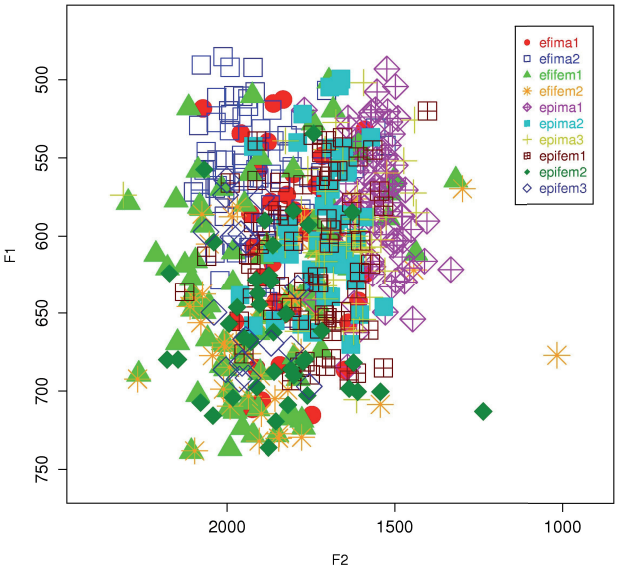


Figure 2 - *All tokens of [ɔ] uttered by all 10 speakers*

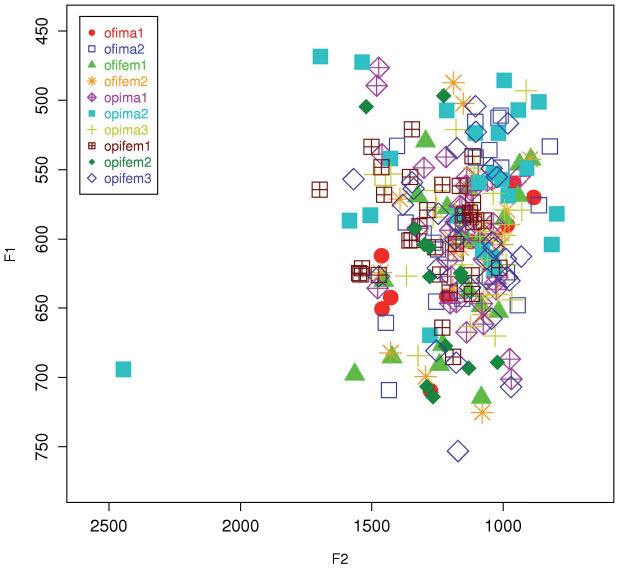
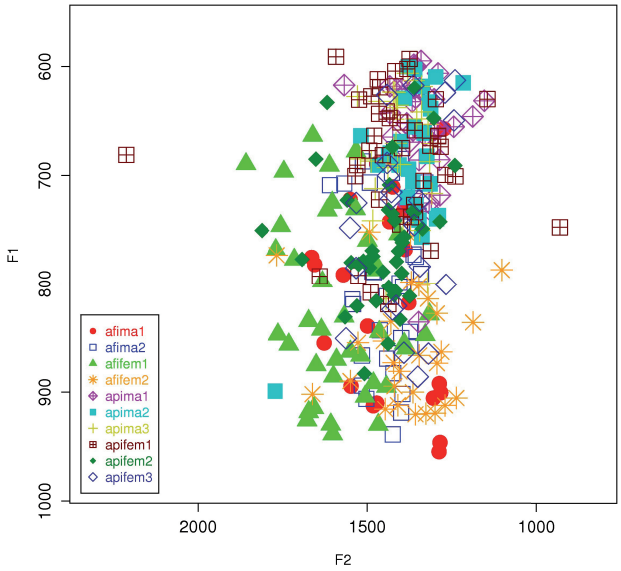
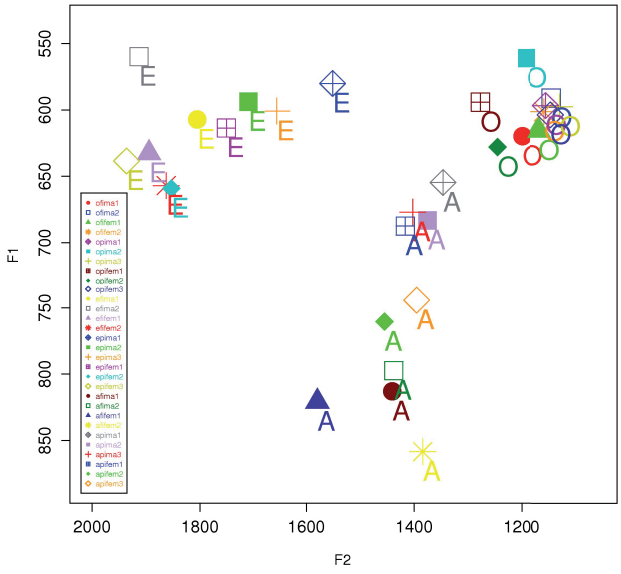


Figure 3 - *All tokens of [a] uttered by all 10 speakers*



The mean values of F1 obtained for the open-mid front unrounded vowel [ɛ] range from 560 Hz to 680 Hz, while the mean F2 values reach values between 1550 Hz and 1900 Hz. The distribution of [ɛ] produced by Pisan male speakers is rather uniform, while there is more variability for their Florentine counterparts. For the female speakers, Pisan and Florentine show similar values.

Figure 4 - *Speakers' means for [ɛ], [ɔ], and [a]*

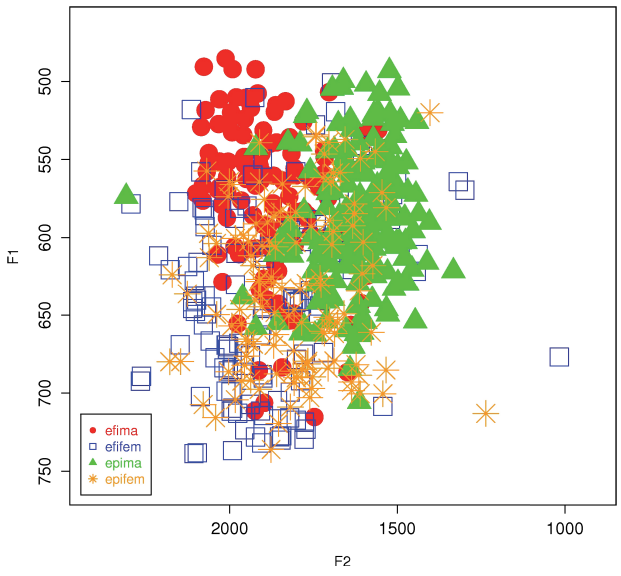


Regarding [ɔ], we noticed that all sets were rather homogeneous, and that the mean values recorded for each speaker and each group were similar, ranging from around 570 Hz to 620 Hz for F1, and from around 1050 Hz to 1250 Hz for F2. The [a] produced by Florentine speakers tended be similar for both males and females (with mean values between 770 Hz to 850 Hz for F1 and 1400 Hz to 1600 Hz for F2). On the other hand, Pisan speakers reveal more variability, especially the female speakers. For example, the average values obtained for one female speaker are comparable to those of the three male speakers. The values for the other two female speakers from Pisa, around 760 Hz for F1 and 1400 Hz for F2, are similar to those produced by their Florentine counterparts. The following graph illustrates the speakers' means for [ɛ], [ɔ], and [a], in support of what has been said in the previous paragraphs (see Figure 4).

The mean values obtained for Pisan [a] (both male and female) tend to be close to the values of [ɔ]. Based on this information, we noticed that perhaps the vicinity between [a] and [ɔ] might not be fortuitous. This could be due to a general shift of the place of articulation of the Pisan vowels. That is, the [ɛ], shifting to lower frequencies, might cause a repositioning of [a], in order to maintain contrast. However, the inconsistent productions for the Florentine variety do not allow us to further explain this behaviour.

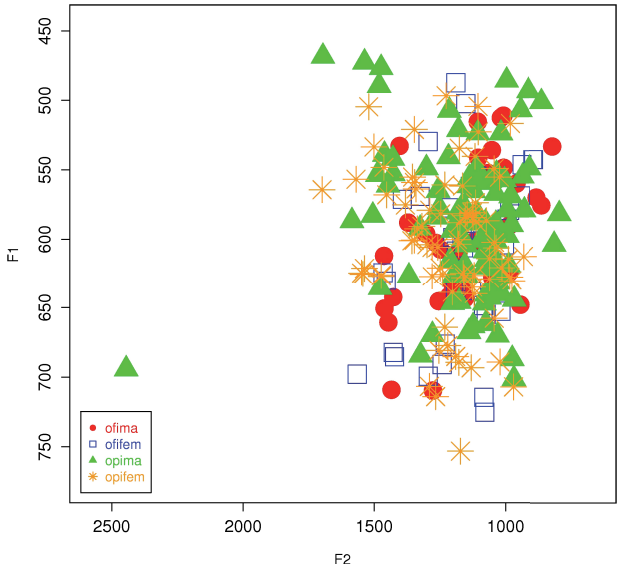
The analysis of the mean values for each speaker suggests that there is a considerable dispersion of the vowels produced by female speakers, both Pisan (epifem) and Florentine (efifem). Plotting as a function of group (Pisan males, Pisan Females, Florentine males, Florentine females), instead of single speakers, provides clearer representations of the vowels' behaviour. It is visible that both male groups (epima and efima) tend to be homogeneous (see Figure 5).

Figure 5 - *Pisan and Florentine groups for [ɛ]*



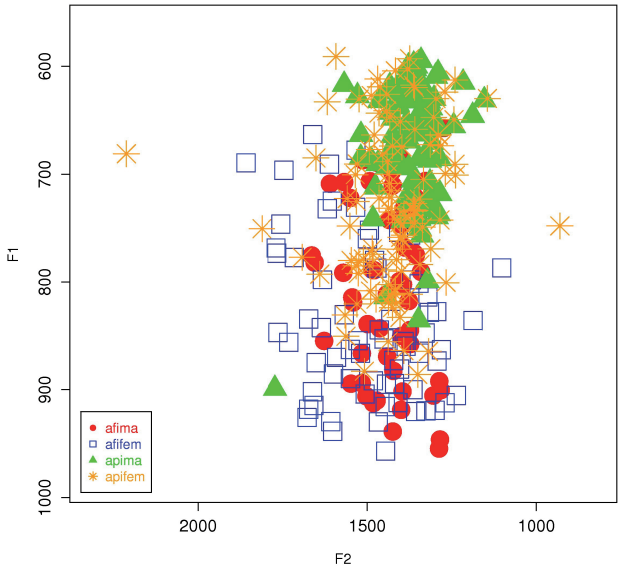
If we consider the graph representing the distribution of [ɔ], it is noticeable that the [ɔ] productions do not allow us to differentiate between the four groups (see Figure 6).

Figure 6 - *Pisan and Florentine groups for [ɔ]*



The tokens of Florentine [a] feature high F1 values. But given the low sample size (4 speakers), it is not possible to establish whether this is a variety-dependent trait.

Figure 7 - *Pisan and Florentine groups for [a]*



The intriguing aspect of this analysis still lies in the vicinity between the values of [ɔ] and [a]. In fact, according to the predictions about the lowering phenomenon, the latter should be closer to [ɛ]. Moreover, the values obtained for [a] display consistent difference only between Pisan males and Florentine males.

So, the results obtained for the four groups were compared with the values reported by Nocchi, Calamai (2009), for [ɛ], [ɔ] and [a] (Table 1). For their research, they analyzed male speakers[11], recorded by means of the map task elicitation technique.

Table 1 - *Nocchi, Calamai (2009) – vowel formants and duration for [ɛ] (Pisan speakers)*

Vowel	*Gender*	*City*	*F1*	*F2*	*F2-F1*
[ɛ]	male	Pisa	538	1693	1155
[ɔ]	male	Pisa	562	1131	569
[a]	male	Pisa	650	1353	703

In the above-mentioned research, the authors described a marked lowering of most of the vowel ellipses and a significant vicinity between [a] and [ɛ]. In fact, it appears that [a] is frequently realized as [æ].

Tables 2, 3 and 4 show the results obtained in our study.

Table 2 - *Mean formant values of [ɛ] as a function of speakers' groups (standard deviations are in brackets)*

[ɛ]		*F1*	*F2*	*F3*	*F2-F1*
Pisa	male	590 (42)	1623 (120)	2625 (138)	1033 (122)
	female	632 (52)	1811 (163)	2779 (300)	1178 (167)
Florence	male	576 (50)	1877 (130)	2673 (150)	1301 (150)
	female	641 (57)	1883 (205)	2746 (298)	1241 (202)

Table 3 - *Mean formant values of [ɔ] as a function of speakers' groups (standard deviations are in brackets)*

[ɔ]		*F1*	*F2*	*F3*	*F2-F1*
Pisa	male	588 (51)	1155 (220)	2654 (201)	566 (229)
	female	604 (53)	1221 (163)	2790 (302)	617 (185)
Florence	male	599 (54)	1150 (131)	2616 (229)	560 (166)
	female	609 (59)	1163 (165)	2624 (232)	554 (155)

[11] Calamai (2004) already presented a thorough review of a cross-gender comparison between Pisan and Leghorn male and female speakers. However, this study analyzed female speakers, employing carrier sentences and controlled speech. Furthermore, it investigated the lowering of open mid-vowels in subjects of different ages.

Table 4 - *Mean formant values of [a] as a function of speakers' groups (standard deviations are in brackets)*

[a]		*F1*	*F2*	*F3*	*F2-F1*
Pisa	male	670 (55)	1372 (81)	2667 (202)	703 (87)
	female	728 (74)	1427 (138)	2669 (296)	699 (150)
Florence	male	804 (82)	1439 (95)	2545 (192)	635 (193)
	female	838 (72)	1498 (160)	2672 (255)	660 (133)

The mean values (Table 2) obtained for the [ɛ] produced by Pisan male speakers (590 Hz for F1, 1623 Hz for F2, 2625 Hz for F3 and 1033 Hz for F2-F1) do not appear to be far from those reported by Nocchi, Calamai (2009) (538 Hz for F1, 1693 Hz for F2, and 1155 Hz for F2-F1), while, on the other hand, the values for Pisan female speakers seem to be vastly different (632 Hz for F1, 1811 Hz for F2, 2778 Hz for F3, and 1178 Hz for F2-F1). However, since this is also the group that displays large internal heterogeneity, it is arguable to what extent we can trust this evidence. As far as [ɔ] is concerned, the means of both Pisan and Florentine groups tend to resemble each other (especially the second and the third formant and the F2-F1 difference), but the standard deviation is rather large. These values are comparable to those found by Nocchi and Calamai (2009). Regarding [a], the mean values for our Pisan male subjects are 670 Hz for F1, 1372 Hz for F2, 2667 Hz for F3 and 703 Hz for F2-F1. These values are perfectly comparable to those found by Nocchi, Calamai (2009) (650 Hz for F1, 1353 Hz for F2, and 703 Hz for F2-F1). Unlike their male counterparts, the Pisan female speakers produced higher mean values (728 Hz for F1, 1427 Hz for F2, 2669 Hz for F3, and 699 Hz for F2-F1), meaning that they are distant from the values presented in Nocchi, Calamai (2009).

4.1 Statistical analysis

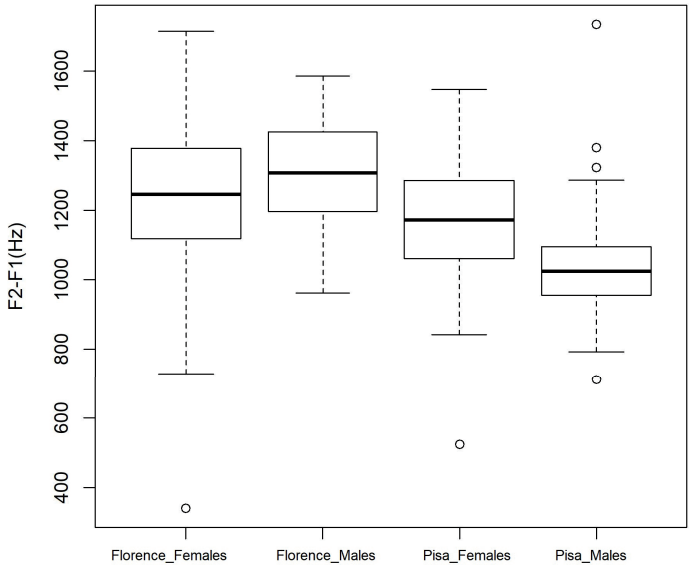
Our *ad hoc* corpus is rather small and not representative of the whole Pisan variety. Indeed, to some degree, it might only be considered representative of the Pisan variety spoken by young educated speakers. Yet, the formant values presented in the previous paragraphs show that there is variability within each group. So, at this point it is questionable whether it is worth carrying out inferential tests. The descriptive statistics (graphs, mean values and standard deviations) introduced earlier might be satisfactory, but a two-way ANOVA with R was also performed to determine whether there was any noteworthy difference between the four groups (Pisan males, Pisan females, Florentine males, and Florentine females).

Based on the results of the pairwise combinations of the continuous variables, and after having analysed the boxplots[12] (see Figure 8, for the mean values of F2-F1 in each of the four groups), we performed an ANOVA using formant values (F1, F2,

[12] Boxplots for [ɔ] and [a] were also analyzed, but for the sake of conciseness this study presents only the [ɛ] boxplot.

and F2-F1) as response variables, and gender and city of origin as factors (categorical variables with two levels each). According to the data presented earlier, the lowering of open [ɛ] and [ɔ] is not significant in the small Pisan corpus investigated here.

Figure 8 - *Boxplots – Pisan and Florentine speakers for [ɛ]*



Since the single and the mean values for [ɔ] were similar across the four groups, and the ANOVA did not show any significant outcomes for this vowel, we are only presenting the ANOVA results for [ɛ][13]. We employed R's "aov" built-in function for the analysis of variance in the corpus, having F2-F1[14] as a response variable and gender and city of origin as factors. The results show no significant main effects for the speakers' gender (Df: 1, F value: 2.822, Pr(>F): 0.1440), but there might be a minor effect, even if not it is not solid, for the speakers' origin (Df: 1, F value: 6.847, Pr(>F): 0.0398) and for the interaction between gender and city of origin (Df: 1, F value: 4.168, Pr(>F): 0.0873). Tukey's honest significant difference (HSD) test allowed us to identify the means that were significantly different from one another. In this test, as far as the speakers' origin was concerned, the means reported for the Pisan and the Florentine groups were significantly different (p adj: 0.0398), while in the case of gender-city interaction the most noteworthy difference was between Pisan male speakers and Florentine male speakers (p adj: 0.0607).

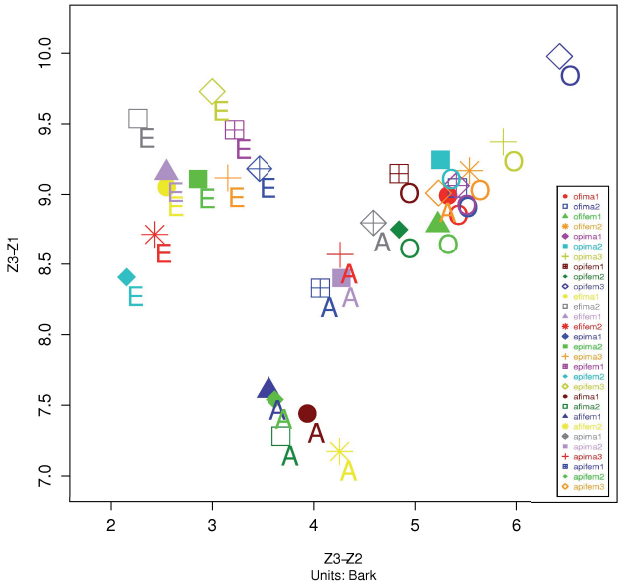
[13] The ANOVA for [a] did not reveal any particular phenomenon.

[14] The outcomes are similar for response variables F1 and F2 alone.

5. *Analysis of normalized vowels*

Since only a few studies aim to combine acoustic and auditory analyses[15], we decided to also include the normalized values of the vowels, following the method used by Ferrero *et al.* (1996). For this reason, we transformed the formant values from Hz to Bark using Traunmüller's Bark conversion formula (Traunmüller, 1990, 1997): $Z_i = 26.81/(1+1960/F_i) - 0.53$, where Z_i is the formant value expressed in Bark, while F_i is the formant frequency in Hz. Then we computed the markers: Z3-Z2, Z3-Z1, and Z2-Z1. The diversity in terms of anatomical features among male and female speakers can then be studied and resized, according to whether the aim is to maximize or minimize the differences between the two systems (Maisano, 1996). So, we performed a Bark Difference Metric Normalization, and then we plotted the vowels using Z3-Z2, Z3-Z1 as coordinates. The vowel space was represented by Z3-Z2 (i.e. Bark-converted F3 minus Bark-converted F2), to model vowel advancement, and Z3-Z1 (i.e. Bark-converted F3 minus Bark-converted F1) to model vowel height. For each speaker, and then for each group, we calculated the mean values and the standard deviation.

Figure 9 - *Speakers' means for [ɛ], [ɔ], and [a]*



Following the procedure shown in the first part of this section, we plotted the mean normalized values for all speakers (see Figure 10, 11, 12). We also computed the mean values and the standard deviation for all groups (see Table 5, 6, 7).

[15] Uguzzoni (1988) and Sorianello (2002) provided auditory classifications, for the local varieties spoken in Pavullo (MO) and Siena, respectively.

Figure 10 - *Pisan and Florentine groups for normalized [ɛ]*

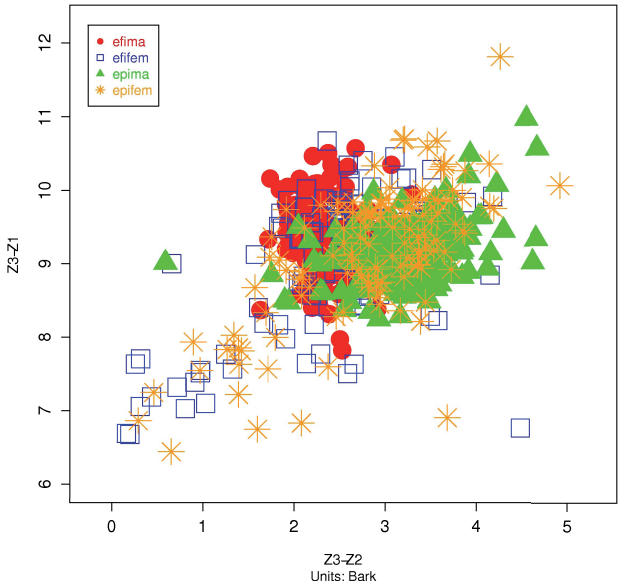


Figure 11 - *Pisan and Florentine groups for normalized [ɔ]*

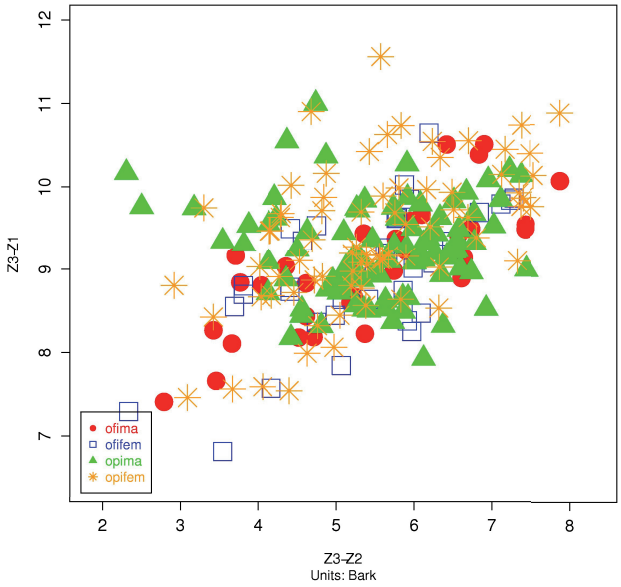
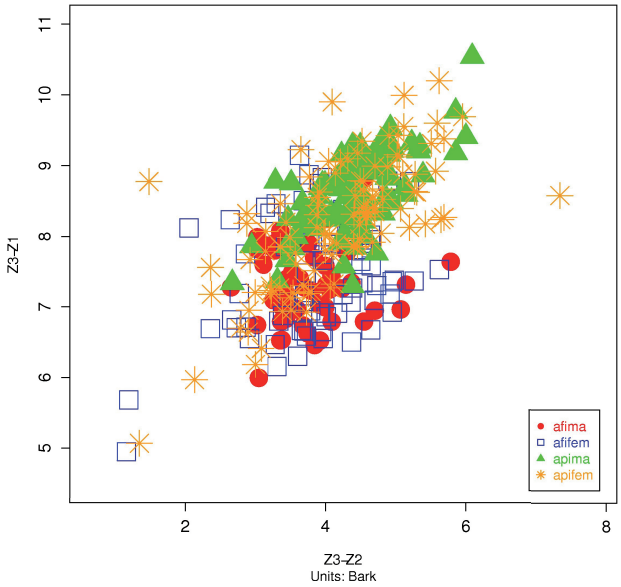


Figure 12 - *Pisan and Florentine groups for normalized [a]*



Tables 5, 6 and 7 show the normalized values of the three vowels obtained for the four groups.

Table 5 - *Mean formant values of [ɛ] as a function of speakers' groups (standard deviations are in brackets)*

[ɛ]		*Z3-Z2*	*Z3-Z1*	*Z2-Z1*
Pisa	male	3.21 (0.56)	9.51 (0.41)	5.93 (0.55)
	female	2.83 (0.84)	9.15 (0.92)	6.32 (0.70)
Florence	male	2.36 (0.34)	9.38 (0.57)	7.02 (0.69)
	female	2.50 (0.85)	9.00 (0.87)	6.49 (0.82)

Table 6 - *Mean formant values of [ɔ] as a function of speakers' groups (standard deviations are in brackets)*

[ɔ]		*Z3-Z2*	*Z3-Z1*	*Z2-Z1*
Pisa	male	5.54 (1.01)	9.22 (0.56)	3.68 (1.16)
	female	5.46 (1.17)	9.40 (0.84)	3.94 (1.01)
Florence	male	5.39 (1.30)	9.03 (0.74)	3.64 (0.86)
	female	5.38 (1.09)	8.97 (0.76)	3.59 (0.83)

Table 7 - *Mean formant values of [a] as a function of speakers' groups (standard deviations are in brackets)*

[a]		*Z3-Z2*	*Z3-Z1*	*Z2-Z1*
Pisa	male	4.40 (0.62)	8.61 (0.56)	4.21 (0.49)
	female	4.14 (0.97)	8.16 (0.90)	4.02 (0.76)
Florence	male	3.79 (0.59)	7.34 (0.57)	3.56 (0.75)
	female	3.85 (0.87)	7.41 (0.82)	3.57 (0.97)

The results confirm that the productions of [ɔ] are similar in all groups, but the standard deviation remains large, since a large number of tokens are distant from the mean. Values for Z3-Z2 range roughly from 5.40 to 5.50 Bark, Z3-Z1 varies between 9.00 and 9.40 Bark, while Z2-Z1 ranges from 3.60 to 3.90 Bark. Regarding [ɛ], the means remain different between the two Pisan groups. Calling to mind the large variability within the female group, the values for Z3-Z2, Z3-Z1, and Z2-Z1 are distributed as follows: 2.83 Bark, 9.15 Bark, and 6.32 Bark for the females, compared to 3.21 Bark, 9.51 Bark, and 5.93 Bark for the males. At the same time, the [ɛ] production by Pisan male speakers is noticeably distinct from the Florentine groups. Finally, the mean values for [a] appear to vary between the four groups, revealing rather substantial differences between Pisan and Florentine male speakers.

5.1 Classification experiment

The intriguing results displayed in the previous sections urged us to test an automatic classification system based on [ɛ]'s specific formant features. The aim was to verify whether this simplified method would allow us to discriminate between the varieties taken into consideration. We performed a series of classification experiments (for gender and city of origin, respectively) with the data mining tool Orange (2013)[16], using the Neural Network technique (sampling: Cross-validation – 10 folds) and the following parameters: Z3-Z2, Z3-Z1, Z2-Z1 expressed in Bark.

Table 8 - *Sampling method and evaluation results for the classification of Pisan [ɛ]*

Groups	*Sampling method*	*Classification method*	*Classification accuracy*
Female_Pisan [ɛ] Male_Pisan [ɛ]	Cross-validation Number of folds: 10	Neural Network	85.94%

[16] Retrieved from https://orange.biolab.si.

Table 9 - *Confusion matrices – Proportions of true and proportions of predicted for the classification of Pisan [ɛ]*

Proportions of true	*Female_Pisan [ɛ]*	*Male_Pisan [ɛ]*
Female_Pisan [ɛ]	75.60%	24.40%
Male_Pisan [ɛ]	6.80%	93.20%

Proportions of predicted	*Female_Pisan [ɛ]*	*Male_Pisan [ɛ]*
Female_Pisan [ɛ]	88.60%	15.50%
Male_Pisan [ɛ]	11.40%	84.50%

Table 10 - *Sampling method and evaluation results for the classification of Florentine and Pisan [ɛ]*

Groups	*Sampling method*	*Classification method*	*Classification accuracy*
Female_Florentine [ɛ] Female_Pisan [ɛ] Male_Florentine [ɛ] Male_Pisan [ɛ]	Cross-validation Number of folds: 10	Neural Network	77.27%

Table 11 - *Confusion matrices – Proportions of true and proportions of predicted for the classification of Florentine and Pisan [ɛ]*

Proportions of true	Female_ Florentine [ɛ]	Female_Pisan [ɛ]	Male_ Florentine [ɛ]	Male_Pisan [ɛ]
Female_Florentine [ɛ]	60.90%	10.20%	16.40%	12.50%
Female_Pisan [ɛ]	3.30%	68.30%	0.00%	28.50%
Male_Florentine [ɛ]	11.40%	4.40%	81.60%	2.60%
Male_Pisan [ɛ]	0.00%	7.40%	0.00%	92.60%

Proportions of predicted	Female_ Florentine [ɛ]	Female_Pisan [ɛ]	Male_ Florentine [ɛ]	Male_Pisan [ɛ]
Female_Florentine [ɛ]	82.20%	11.30%	18.40%	7.40%
Female_Pisan [ɛ]	4.20%	73.00%	0.00%	16.10%
Male_Florentine [ɛ]	13.70%	4.30%	81.60%	1.40%
Male_Pisan [ɛ]	0.00%	11.30%	0.00%	75.10%

The evaluation results for [ɛ] showed that this minimalist Neural Networks classification system based only on Z3-Z2, Z3-Z1, and Z2-Z1 values can predict with reasonable accuracy (over 85%) whether the vowels were uttered by Pisan male or Pisan female speakers.

If we extend this investigation by adding the two Florentine groups the classification accuracy is slightly lower (77%), but still promising. This evidence supports

the previous impressions, namely that, for [ɛ] productions, the most significant difference is between the two male groups. Nevertheless, due to the limited number of productions analyzed here, further investigations should be carried out.

6. *Conclusions and further research*

This pilot study aimed to detect and describe gender-specific variation in vernacular speech, by examining the lowering of open-mid vowels in a group of young Pisan speakers. In general, the data did not reveal commensurable differences in terms of vowel quality ascribable to the lowering phenomenon. In regards to gender differences, the Pisan and Florentine female groups showed more internal variation, compared to their male counterparts. On the other hand, the lowering of [ɛ] detected in Pisan male speakers appears clear when investigating the F2-F1 difference. However, the numerous creaky voice productions of the two Pisan male subjects (Calamai, 2015; Melvin, Clopper, 2015) in some cases hindered accurate measurements of the vowel quality of this group[17].

In order to discern these preliminary results, we carried out an automatic classification experiment based on [ɛ]'s specific formant features. The results indicated that it is possible to differentiate between [ɛ] as uttered by the four groups, but the experimental conditions (i.e. limited number of speakers and the type of system employed) are not sufficient to speculate on the reasons behind these differences. Based on this classification experiment, we can only assert that the Bark-converted values Z3-Z2, Z3-Z1 and Z2-Z1 could represent fairly robust features for automatic vowel discrimination.

On a different note, we believe that the level of education and social class (Nodari, 2016) might have a greater effect than gender on discriminating between speakers. In particular, one could assess the role of education in triggering or hindering the lowering of [ɛ] and [ɔ] in the Western area of Tuscany. Arguably, the lowering of open-mid vowels was not attested in our corpus because all speakers belong to the same social network, have comparable family compositions (upper-middle class) and the same level of education (five of them are enrolled at the University). For these reasons, the study should be extended to a larger and more wide-ranging group of subjects in order to obtain more reliable results. Finally, vowel duration and stress could also be considered in future sociophonetic investigations.

Acknowledgments

We would like to thank Prof. Giovanna Marotta for the constructive suggestions and the supervision of our work. Her recommendations have been helpful for the interpretation of the overall vocalic shift that might occur in the Pisan variety. Her proposal to use education as a discriminant sociolinguistic factor has also opened new dimensions for this research in the future.

[17] The use of creaky voice is also highlighted in the literature (Calamai, 2015; Sorianello, 2006) as a typical male feature in low registers of both Tuscan and Roman varieties.

Bibliography

Adank, P., Smits, R. & Van Hout, R. (2004). A Comparison of Vowel Normalization Procedures for Language Variation Research. In *Journal of the Acoustic Society of America*, 116(5), 3099-3107.

Albano Leoni, F. (coordinator) (1999). *Archivio del Parlato Italiano (API)*. http://www.parlaritaliano.it/index.php/it/dati/40-api-archivio-del-parlato-italiano.

Albano Leoni, F. (2007). CLIPS – Corpora e Lessici di Italiano Parlato e Scritto. In *Linguistica computazionale: ricerche monolingui e multilingui*. http://www.clips.unina.it/it/corpus.jsp/.

Berruto, G. (2005). *Fondamenti di sociolinguistica*. Bari: Laterza.

Boersma, P., Weenink, D. (2016). *Praat: doing phonetics by computer.* http://www.praat.org/.

Byrd, D. (1994). Relations of sex and dialect to reduction. In *Speech communication*, 15, 39-54.

Calamai, S. (2001). Il vocalismo atono della varietà pisana. Prime evidenze sperimentali. In Regnicoli, A. (Ed.), *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*. Roma: Il Calamo.

Calamai, S. (2003). Vocali fiorentine e vocali pisane a confronto. In Albano Leoni, F., Cutugno, F., Pettorino, M., Savy, R. (a cura di), *Atti del Convegno Nazionale Il parlato Italiano*. Napoli; D'Auria, 1-25.

Calamai, S. (2004). *Il vocalismo tonico dell'area pisana e livornese. Aspetti storici, percettivi, acustici*. Alessandria: Edizioni dell'Orso.

Calamai, S. (2015). *Introduzione alla sociofonetica*. Roma: Carocci.

Calamai, S., Ricci, I. (2005). Sulla percezione dei confini vocalici in Toscana: primi risultati. In Cosi, P. (Ed.), *Proceedings of the 1st AISV Conference*, Torriana: EDK Editore, 63-87.

Chambers, J.K. (2002). Patterns of variation including change. In Chambers, J.K., Trudgill, P., Schilling-Estes, N. (Eds.), *The Handbook of Language Variation and Change*. Oxford: Blackwell Publishing.

Cosi, P., Ferrero, F. & Vagges, K. (1995). Rappresentazioni acustiche e uditive delle vocali italiane. In *Proceedings of the 23rd National Conference of Associazione Italiana di Acustica*, Bologna, 12-14 Settembre 1995, 151-156.

Eckert, P. (1989). The whole woman: sex and gender differences in variation. In *Language variation and change*, 1(3), 245-267.

Elan. https://tla.mpi.nl/tools/tla-tools/elan/. Nijmegen: Max Planck Institute for Psycholinguistics.

Fabricius, A. (2008). *Vowel Normalization in Sociophonetics: When, Why, How?* Seminar for Sociolinguistics Circle. Copenhagen University.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Ferrero, F.E., Magno Caldognetto, E. & Cosi, P. (1996). Sui piani formantici acustici e uditivi delle vocali di uomo, donna, bambino. In Peretti A., Simonetti P. (Eds.), *Proceedings of the 24th National Conference of Associazione Italiana di Acustica*. Padova: Arti Grafiche Padovane, 169-178.

Filipponio, L., Cazzorla, S. (2016). The vowels of Bari. A comparison between local dialect and regional Italian. In Vayra, M., Avesani, C. & Tamburini, F. (Eds.), *Il farsi e il disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio.* Milano: Officinaventuno, 59-71.

Giannelli, L. (2000). *Toscana*. Pisa: Pacini.

Kendall, T., Thomas, E. (2009). *Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package*, https://cran.r-project.org/package=vowels.

Klatt, D.H., Klatt, L.C. (1990). Analysis, synthesis and perception of voice quality variations among female and male talkers. In *Journal of the Acoustic Society of America*, 87(2), 820-857.

Maisano, L. (1996). Sistemi vocalici maschili e femminili a confronto. In Peretti, A., Simonetti, P. (Eds.), *Proceedings of the 24th National Conference of Associazione Italiana di Acustica.* Padova: Arti Grafiche Padovane.

Marotta, G., Calamai, S. & Sardelli, E. (2004). Non di sola lunghezza. La modulazione di f0 come indice socio-fonetico. In De Dominicis, A., Mori, L. & Stefani, M. (Eds.), *Costituzione, gestione e restauro di corpora vocali*. Roma: Esagrafica, 210-215.

Melvin, S., Clopper, C. (2015). Gender Variation in Creaky Voice and Fundamental Frequency. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: The University of Glasgow. Paper number 320.

Nocchi, N., Calamai, S. (2009). Durata e strutture formantiche nel parlato toscano: Indagini preliminari su un campione di dialoghi semispontanei. In Schmid, S., Schwarzenbach, M. & Studer, D. (Eds.), *Proceedings of the 5th Conference of Associazione Italiana di Scienze della Voce. La dimensione temporale del parlato.* Torriana: EDK, 195.

Nodari, R. (2016). Descrizione acustica delle occlusive sorde aspirate: analisi sociofonetica dell'italiano regionale di adolescenti calabresi. In Vayra, M., Avesani, C. & Tamburini, F. (Eds.), *Il farsi e il disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio.* Milano: Officinaventuno, 139-153.

Orange (2013). *Data Mining Toolbox in Python*. https://orange.biolab.si.

R Development Core Team (2008). *R: A language and environment for statistical computing*. https://www.r-project.org. Wien: R Foundation for Statistical Computing.

Romito, L., Ciardullo, M.A. & Tarasi, A. (2016). Analisi acustica delle occlusive sorde aspirate del dialetto di San Giovanni in Fiore (CS). In Vayra, M., Avesani, C. & Tamburini, F. (Eds.), *Il farsi e il disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio*. Milano: Officinaventuno, 169-186.

struttura sonora del linguaggio. Milano: Officinaventuno, 169-186.

Sorianello, P. (2002). Il vocalismo dell'italiano senese: un'indagine sperimentale. In Regnicoli, A. (Ed.), *La fonetica acustica come strumento di analisi della variazione linguistica in Italia.* Roma: Il Calamo, 47-52.

Sorianello, P. (2006). Per una rappresentazione uditiva dei segmenti vocalici: il caso del senese. In Agonigi, M., Cioni, L. & Paradisi, E. (a cura di), *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa*, Nuova Serie, 159-173.

Syrdal, A.K., Gopal, H.S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. In *Journal of the Acoustical Society of America*, 79(4), 1086-1100.

Tamburini, F. (Eds.), *Il farsi e il disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio.* Milano: Officinaventuno, 169-186.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. In *Journal of the Acoustic Society of Am*erica, 88(1), 97-100.

Traunmüller, H. (1997). *Auditory scales of frequency representation*. https://www.ling.su.se/staff/hartmut/bark.htm.

Uguzzoni, A. (1988). Verso un'analisi parametrica del vocalismo di un dialetto frignanese. In *Italia Dialettale*, 25, 86-110.