

Spectral characteristics of voiceless fricative consonants in Florentine Italian

Patrizia Sorianello

Scuola Normale Superiore di Pisa, Italy

E-mail: sorianello@libero.it

ABSTRACT

Tuscan Italian is characterized by the spirantization of voiceless stops. This process has been well studied in the past from different points of view. However, instrumental phonetics surveys are, at present, very few. The purpose of the research is to examine the spectro-acoustic properties of Florentine fricative allophones. In particular, the noise of voiceless fricatives is investigated using FFT spectra taken from extensive corpora of quasi-spontaneous dialogues produced by various speakers. The analysis seems to suggest the presence of a continuum of allophonic realizations. The location of energy in the spectrum plays a relevant role in the place of articulation detection of these fricative sounds. The analysis performed has allowed to establish with more precision the phonetic nature of the allophonic variants.

1. INTRODUCTION

A process of consonant weakening, best known as spirantization or Tuscan ‘gorgia’, is typical of the Italian spoken in Tuscany. Both voiceless and, in a more variable degree, voiced stops are the target of the process. The most common outcome is a fricative sound, i.e. /p t k/ > [ɸ θ h/ɦ], /b d g/ > [β ð ɣ]. This phonological phenomenon occurs in a postvocalic context; more exactly, the stop must be preceded by a vowel and can be followed by a vowel or glide or liquid segment; for instance, *manica* (‘sleeve’) [ˈma:niha], *abete* (‘fir-tree’) [aˈβe: θe], *capire* (to understand) [kaˈɸi:re], but also *la pietra* (‘the stone’), [laˈɸjɛ: θra], *lacrima* (‘tear’) [ˈla:hrima].

The process does not involve the three stops in a similar way, since there is a sort of correlation between the place of articulation of the stop phoneme and the degree of spirantization. The velar obstruent /k/ is the primary target of the ‘gorgia’, progressively followed by /t/ and /p/. The spirantization is very widespread, nevertheless it has not a regular distribution within the Tuscan region; in the central area (Florence, Siene and Pistoia), the spirantization shows its greatest diffusion. This topic has been widely investigated in the past; several theoretical frameworks have been taken into account (cf., at least, [1], [2], [3], [4], [5]).

However, there is a substantial lack of acoustic studies on this specific subject. The spectro-acoustic properties of

these allophones have never been described. Up to now, the most careful phonetic classification of the possible sounds produced by the Tuscan spirantization is the one reported in [2] and [3]. The authors recognize three different degrees of oral opening in the production of bilabial, dental and velar obstruents, but two degrees only for the glottal place of articulation. Recently, the spirantization of Pisa Italian has been studied from an acoustic point of view (see [6], [8]). In these experimental analyses, the segmental duration of VOT was considered as a basic parameter for the discrimination of the allophones. A remarkable outcome of these analyses concerns the identification of voiceless stops with a long VOT, a kind of segment generally not included in the literature regarding the ‘gorgia’.

The present research aims at analyzing the spectral features of the fricatives produced as allophones of the voiceless stop phonemes in Florentine Italian. The results of a detailed acoustic analysis are reported and discussed. The fricatives found are classified according to their articulatory strength. A new phonetic transcription is also proposed in order to better keep distinct similar fricative segments.

2. METHOD

Two informal dialogues produced by three young non-professional native speakers of Florentine Italian were analyzed. The subjects were two males (P1, P3) and a female (P2). The corpus was formed by 500 voiceless stops including 150 bilabial stops, 180 dental stops and 170 velar ones. Only intervocalic segments were considered either across or within words; they are followed by any vowel and without any distinction between stressed and unstressed syllables.

Speech material was digitized at 22050 Hz sampling rate and 16-bit quantization. The acoustic analysis was performed by means of *Kay MultiSpeech* software (*Model 3700*, version 2.2). For each stop the following acoustic parameters were measured:

- 1) segment duration,
- 2) frequency range of the fricative noise,
- 3) frequency position of the most prominent spectral peak of the consonant,
- 4) intensity value of the next vowel,
- 5) amplitude difference between the consonant and the following vowel.

The amplitude and the frequency of the frication noise were obtained by a *FFT Power Spectrum* extracted at cursor in correspondence of the most intense point

identified in the spectrogram (points 128, 6 ms window). As far as the vowel is concerned, the first three contiguous glottal pulses were considered for the extraction of an average FFT spectrum (*Long Term Average Power Spectrum*).

3. RESULTS

In the intervocalic context we observed a continuum of allophonic realizations comprehending voiceless, lenis and voiced stops, as well as voiceless, voiced fricatives and approximants. During the analysis we revealed a certain degree of speaker-based variation. This result is consistent with previous research (cf. [8]). More precisely, the male speaker P1 presents both voiceless and voiced fricatives for the bilabial and dental segments, while the male speaker P3 shows no voicing at all for these same segments.

All the three subjects alternate two voiceless spirants allophones. The first has a strong and compact fricative noise, whereas the second displays a weaker and very diffuse noise. In order to phonetically separate these sounds, a different transcription was employed. We used the IPA diacritic [̥] for the first category of sounds, i. e. [̥ɸ ̥θ ̥h]. With regard to fricatives, this IPA diacritic represents a closer or higher variant of a same consonant, i.e. a fricative with a narrower closure of the vocal tract than IPA [ɸ θ h]. We also distinguished two different allophones within the voiced series: the former one has a diffuse fricative noise, while the latter one exhibits an evident formant structure. The phonetic transcription employed for these two groups of allophones, respectively [β δ γ] and [β̥ δ̥ γ̥], was differentiated by means of the IPA diacritic [̥], a symbol indicating the lowering of the place of articulation and, therefore, a more open sound. This IPA diacritic allows us to disambiguate the two possible phonetic outcomes of /b d g/, namely voiced fricatives or frictionless approximants.

Table 1 displays the percent values of occurrence of each Florentine allophone. The fricatives are very frequent, especially for bilabial (58.4%) and dental places of articulation (86.2%). For /k/ the most common phonetic outcome is the approximant [ɦ] (45.1%), followed by the voiceless glottal fricative [h] (34%). The segment /k/ is totally deleted in the 13% of the total recurrences. For the bilabial phoneme, we also found a remarkable occurrence of a lenis stop (24%), an allophone never attested for dental and velar places of articulation up to now. By lenis stop we mean a segment without burst and VOT, often with a weak voice bar and some traces of fricative noise. This sound represents the first reduction level of a voiceless stop: the lack of burst denotes a consonant with a weaker articulation (cf. also [9]). The highest percent value of voiceless stops was observed for /p/; this is the proof that the bilabial stop is less sensible to Tuscan spirantization (cf. also [6]). In the whole corpus the occurrence of a voiceless stop is context-dependent: it occurs only in hyperarticulated or emphasized words.

The production of a voiceless phoneme as voiced segment must be considered as a sign of articulatory weakness. We found voiced stops [b d g], voiced fricatives

[β δ γ] and finally approximant sounds [β̥ δ̥ γ̥]. By approximant, we mean a segment which does not involve any kind of closure of the vocal tract. As expected, the glottal approximant is very frequent for intervocalic /k/; on the opposite the occurrence of [β] is scarce, only 1.1%.

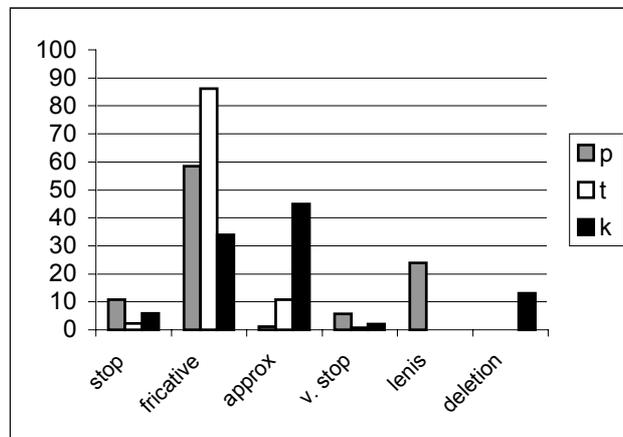


Table 1: percentage of the Florentine allophones of /p t k/: (voiceless stops, fricatives, approximants, voiced stops, segment deletion).

Spectral data: The results obtained indicate that fricative allophones are distinguished from each other on the ground of a number of different acoustic cues. Each allophone shows a distinct spectral pattern, as represented in Tables 2-4 where we report the frequency value of the consonant, its duration, the intensity value for the consonant and the following vowel and, finally, the ratio obtained subtracting the intensity of the vowel from that of the preceding consonant. Spectral results demonstrate that the relevant acoustic cue for both [̥ɸ] and [̥β] lies in a diffuse spectrum. Normally, these sounds do not show a clear area of spectral prominence, since their spectrum is rather flat with a very weak noise component (cf. Fig. 1). However, the frequency location of the main prominent spectral peak is 3832 Hz on average for the voiceless segment and 4791 Hz for the voiced one. Consequently, their intensity values is low too, exactly 22 dB for [̥ɸ] and 21 dB for [̥β]. In a similar way, the energy of [̥θ] is distributed throughout the spectrum, the region in which there is a greater intensity is between 4500/5000 Hz (cf. Fig. 2). The spectral shape of [̥θ] is very different, since the sound does not show energy below 4000 Hz, as the major noise concentration has been observed around 5000 Hz up to cover the highest frequencies (cf. Fig. 3). The most intense peak is located at 5721 Hz. In addition, [̥θ] always reveals a greater amplitude value compared to [̥θ], respectively 40 dB and 29 dB. The frequency location of the spectral peak progressively lets down in [̥δ] and [̥γ], as well as their intensity values and segmental duration.

As far as /k/ is concerned, we had to set apart the adjacent vowels into two groups with reference to the value of the front feature. As is widely known, /k/ has two consonantal *loci*, according to the quality of the next vowel. If the Florentine segment is followed by a front vowel, the spectral energy is always higher and more intense than when the same allophone is followed by a back vowel. In

this latter case, the prominence of [h] is only 868 Hz (see Table 4 and Fig. 6). Some coarticulatory influences are acoustically present in the fricative sound. Taking into account these specific spectral data, we may assume that in a back vocalic environment, the fricative could be better classified as [x] rather than as [h]. It's worth remembering that [x] shows a very low frequency noise, generally coincident with the F2 of the following vowel. The velar fricative, already included in the classification provided by [3] and [4], seems to be the normal outcome for intervocalic /k/ in Pisa Italian, regardless of the vowel type (cf. [6], [7]). Unlike what happens in Pisan speech, in Florentine Italian the occurrence of this segment appears to be strongly conditioned by coarticulation.

We identified two glottal fricatives, respectively [h] and [ɦ] (cf. Figg. 4-5). The first exhibits a very diffuse fricative noise with an average intensity of 30 dB, while [ɦ] shows a clear formant structure, whose amplitude is lower with respect to that of the adjacent vowels.

Duration: In the data collected, allophones and temporal duration seem to be strictly related: the weaker the consonant, the shorter the segmental duration. Fricative segments are generally shorter than stops, approximants are shorter than fricatives. More in detail, the longest allophones are always the voiceless stops, on average [p]= 70 ms, [t]= 67 ms, [k]= 88 ms. They are progressively followed by the voiceless fricatives, [ɸ]= 51 ms, [θ]= 66 ms, [h]= 63 ms; then we find the voiced fricatives [β]= 51 ms, [ð]= 40 ms [ɦ]= 44 ms; finally the approximants, [ç]= 28 ms, [ɣ]=39 ms. The production of approximants sounds can be regarded as the endpoint of this gradient weakening process. The next possible reduction is the total deletion of the segment, as frequently recorded for /k/, i. e. *amico* ('friend') [a'mi:ɔ].

Consonant intensity: The weakest voiceless consonant among all is the bilabial [ɸ]: it is -7 dB with respect to [θ], -18 dB to [ç] and -9 dB to [h]. On the contrary, the strongest fricative is [ç] which differs from [θ], since this latter has always a lower amplitude, although, as stated above, the main difference, lies in the location of fricative noise within the frequency range. As expected, voiced fricatives show a decrease intensity, in comparison to the correspondent voiceless allophone. This is true for bilabial and dental places of articulation, but not for the velar one. We found the following data: [β] vs. [ɸ] = -2 dB, [ð] vs. [θ] and [ç] respectively - 4 and - 15 dB. [ɦ] and [h] have a similar intensity, but their spectrographic structure is very different, as [h] presents a friction noise, whereas [ɦ] shows a formant structure. [ɦ] must be regarded as an approximant sound rather than a voiced fricative. In this respect, the behaviour of the other approximants is very interesting: [ç] is more intense of [ð]; in a same way, [ɣ] has a higher amplitude of both [h] and [ɦ]. Thus, an increase in intensity seems to correspond to a weakening of the consonant. The

data are in line with the experimental findings reported in [9] for English and Spanish.

/p/	C-FFT Hz	C ms	VOT ms	C-FFT dB	V-FFT dB	C-V dB
p	2274 (401)	70 (8)	21 (9)	24(6)	52 (5)	-28
ɸ	3831 (863)	51 (16)	--	22 (6)	54 (7)	-32
β	4791 (988)	51 (16)	--	20 (4)	52 (2)	-32
b	--	59 (11)	--	--	54 (3)	---
b	--	61 (10)	--	--	46 (3)	---

/t/	C-FFT Hz	C ms	VOT ms	C-FFT dB	V-FFT dB	C-V dB
t	4045 (901)	67 (23)	25 (5)	34 (2)	40 (3)	-6
ç	5721 (602)	73 (14)	--	40 (5)	49 (6)	-9
θ	5317 (609)	66 (17)	--	29 (5)	48 (6)	-19
ð	3663 (133)	40 (15)	--	25 (2)	46 (4)	-21
ç	962 (66)	28 (3)	--	31 (6)	48 (4)	-17
ɦ	2525 (899)	37 (8)	--	29 (5)	54 (5)	-25

/k/	C-FFT Hz	C ms	VOT ms	C-FFT dB	V-FFT dB	C-V dB
k	2355 (612)	88 (11)	35 (5)	32 (6)	55 (3)	-23
h+V [+front]	2838 (441)	58 (9)	--	32 (4)	55 (4)	-23
h+V [-front]	867 (182)	69 (19)	--	30 (7)	55 (3)	-25
ɦ+V [+front]	3424 (872)	40 (9)	--	32 (6)	48 (8)	-16
ɦ+V [-front]	2047 (982)	47 (7)	--	29 (4)	53 (6)	-24
ɣ	515 (221)	39 (9)	--	41 (3)	51 (2)	-10

Tables 2/4: Mean values and standard deviation of: spectral frequency for the allophones of /p/, /t/ /k/, consonant duration., VOT, intensity value for the consonant and the following vowel, intensity ratios.

Ratio C-V= the best way of evaluating the intensity value of a consonant is establishing a comparison with another sound of the same segmental string: intensity is in fact only a relative acoustic parameter. We, therefore, calculated a ratio subtracting the intensity value of the vowel from that of the preceding consonant. In our data, this ratio is always expressed by a negative number, that means that the consonant has never a greater amplitude than the vocalic segment. However, the highest difference is observed for [ɸ], (-32 dB). This *datum* is not surprising, since the segment is the weakest fricative. On the contrary, the smallest difference is found for [ç], (-9 dB), which is the strongest fricative among those occurring in our corpus. An intermediate value between bilabial and interdental fricatives, i. e. - 22 dB on average, is indeed the difference calculated for [h]/[ɦ].

With respect to the voiced segments, we found a ratio of -32 dB for [β], -21 dB for [ð] and -20 dB for [ɦ]. Consonants examined show different degrees of reduction. Normally, a weaker segment is also more intense, since it increases in sonority becoming more similar to a vowel. This explains why the voiced velar approximant [ɣ] is more intense of both [h] and [ɦ], the ratio C-V is only -10 dB, and why [ç] has a greater amplitude than [ð] and [θ], but

not than [θ] and [t].

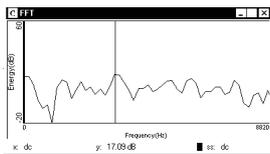


Fig. 1: FFT spectral pattern of [ϕ]

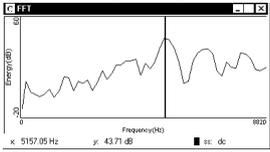


Fig. 2: FFT spectral pattern of [θ]

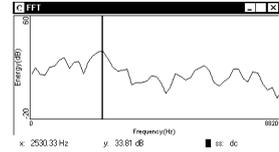


Fig. 3: FFT spectral pattern of [h]

Fig. 4: FFT spectral pattern of [h]

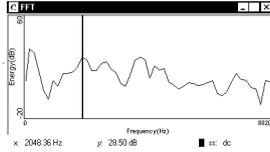


Fig. 5: FFT spectral pattern of [ɦ]

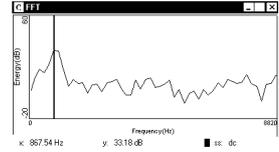


Fig. 6: FFT spectral pattern of [x]

4. DISCUSSION AND CONCLUSION

Our acoustic analysis confirms that Florentine ‘gorgia’ is not a mere spirantization process, but rather a phenomenon of continuous consonant reduction. The gradient nature of the process determines different degrees of consonant weakening: lenition, voicing as well as spirantization, approximantization and deletion. In the process, the place of articulation plays a relevant role. Both approximantization and deletion are common for the intervocalic /k/, whereas they do not affect /p/. Therefore, the bilabial consonant shows a stronger resistance degree as compared to /t/ and /k/. Each stadium of reduction is marked by a different value of duration and intensity. In addition, spectral patterns have a significant weight in the detection of the place of articulation. On one hand, the frequency distribution of the noise is tightly dependent on the place of articulation of the fricatives. On the other one, the degree of supraglottal constriction, the shape of oral tract, as well as the pressure and the rate of airflow are all articulatory cues conditioning the intensity of the noise.

We would like to point out that the acoustic properties of Italian fricatives have not been studied up to now. Therefore, we had to compare our data with those available in the literature regarding other languages. Our data are consistent with previous results about fricatives (cf. among all [10], [11], [12]). As expected, in Florentine Italian, the voiceless bilabial fricative has a low intensity noise and a very long spectrum, as recorded for other languages. As it is well known, there is a positive correlation between the length of the front cavity and the resonant frequencies of the sound: the longer the cavity the lower the frequencies. The spectral properties observed for [h] and [ɦ] agree too with the trends already noted in past research. However, the Florentine interdental fricative seems to be different from the English [θ] (cf. [13]), since the latter consonant shows a flat spectrum stretching over the whole frequency range with a peak centered about 8000 Hz, whereas in the former fricative the noise is very often located in a well definite

frequency region.

In summary, we retain that our results provide useful spectral information about Tuscan fricative allophones. At the same time, we are perfectly aware that further experimental analysis is needed in order to achieve an extensive acoustic picture of Tuscan spirantization.

REFERENCES

- [1] G. Contini, “Per una interpretazione strutturale della cosiddetta ‘gorgia toscana’ in *Boletim de Filologia* vol. 19, pp. 269-281, 1960.
- [2] L. Giannelli, *Toscana, Profilo dei Dialetti Italiani*, Pisa: Pacini, 1976.
- [3] L. Giannelli L. & L.M. Savoia, L’indebolimento consonantico in Toscana, I, *Rivista Italiana di Dialettologia* vol. 2, , pp. 25-58; 1978.
- [4] L. Giannelli & L.M. Savoia, L’indebolimento consonantico in Toscana, II, *Rivista Italiana di Dialettologia* 3-4, pp. 39-101, 1979-1980
- [5] L. Bafille, La spirantizzazione toscana nell’ambito della teoria degli elementi, in AA.VV., *Studi linguistici offerti a G. Giacomelli dagli amici e dagli allievi*, Padova: Unipress, pp. 27-38, 1997.
- [6] G. Marotta, E. Bertoni, R. Franceschi, M.F. Giuliani, P. Soriano, “Le occlusive sorde dell’italiano parlato a Pisa:varianti aspirate e fricative”, *Proceedings of the XII Giornate di Studio del G.F.S.*, Macerata 2001.
- [7] G. Marotta, “La gorgia rivisitata acusticamente e ottimalmente” *Proceedings of the National Congress Il parlato Italiano*, Naples 2003, in press.
- [8] G.W Hughes & Halle M., “Spectral properties of fricative consonants”, *JASA* 28, 303-310, 1956.
- [9] L. M. Lavoie, *Consonant strength*, New York: Garland Publishing, 2001.
- [10] P. Strevens, “Spectra of fricative noise in human speech”, *Language and Speech* 3, pp. 32-49, 1960.
- [11] W. Jassem, “Acoustical description of voiceless fricatives in terms of spectral parameters”, *Speech Analysis and Synthesis* 1, pp. 189-206, 1968.
- [12] S. J. Behrens & S. E. Blumstein, “Acoustic characteristics of English voiceless fricatives: a descriptive analysis”, *Journal of Phonetics* 16, pp. 295-298, 1988.
- [13] M.Tabain, “Non-sibilant fricatives in English: spectral information above 10Khz”, *Phonetica* 55, pp. 107-130, 1998.