web-sls 98.01 20.07.98

WEB-SLS

The European Student Journal of Language and Speech

Acoustic Analysis of Singleton and Geminate Fricatives in Italian

Maurizio Giovanardi and Maria-Gabriella Di Benedetto

INFOCOM Department

University of Rome - La Sapienza

Via Eudossiana, 18

IT-00184 Roma

gaby@acts.ing.uniroma1.it

Abstract

Acoustic correlates of singleton *vs*. geminate fricatives in Italian were investigated on the basis of acoustic analyses carried out on Italian disyllabic words. In the analyzed words, the singleton *vs*. geminate consonant appeared in the symmetrical context of the three Italian vowels [a, i, u]. Time related and frequency related parameters were examined. Time parameters were all based on durational measurements performed within the consonant and surrounding vowels. Frequency parameters, such as formants, fundamental frequency, and energy based parameters were computed at different sampling points all through the analyzed words. Results showed that the duration of the consonant was significantly different in singleton *vs*. geminate words, as well as the duration of the vowel preceding the consonant, while the other analyzed parameters were not. These results are in agreement with previous studies reported in the literature on the gemination of gemination of Italian stop consonants. In particular, it was observed that the ratio between the durations of the consonant and pre-consonant vowel was significantly related to gemination.. It was found that a value of about 1 of the above ratio discriminated singleton *vs*. geminate fricatives. In addition, this value was verified to be significant also for stops. This result leads to a suggestion that, possibly, speaker intention in producing a geminate is reflected in the production of a consonant at least longer than the vowel preceding it.

top of article

Introduction

Gemination can be defined as the clustering of a single consonant into a 'double' or geminate consonant. This phenomenon plays a major role in Italian, a language in which several words change their meaning as a function of the presence or absence of gemination of one consonant in the word. Most often in Italian these words are disyllabic, with the stress placed on the first syllable of the word. However, gemination can also be observed across words of a same sentence.

Only a few studies on gemination are reported in the literature. Recent papers on stop consonants report that there is "a strong correlation between the presence of gemination and the first vowel and occlusive silence duration" (Rossetti 1993/94 [2], [3]). This effect was found to be significant also in Hindi geminate consonants (Shrotriya *et al*. 1995 [4]). Moreover, it has been reported that the duration of the utterances, singleton or geminate roughly remains constant. In particular "that speakers unconsciously tend to maintain, by balancing the durational change of some of the phonetic segments with the durational change of others, the rhythmic structure of a word" (Esposito, Di Benedetto 1997 [5]). This idea is also supported by Blumstein *et al*. 1998 [6], in which it is affirmed that "the ratio between consonant duration and preceding vowel duration discriminated between singletons and geminates both within and across speaking rate".

The present study analyzes the phenomenon of gemination in Italian fricatives. Since the papers mentioned above only dealt with stop consonants, our aim was also to find out similarities and differences in the case of fricatives.

The set of fricatives which can be geminated in Italian, within words, are [f, v, s], which represents a subset of all Italian fricatives. To avoid any influence from the meaning of the words on speakers, the above consonants were included in disyllabic words with no real Italian sense. In fact, all words considered were structured as [vowel-consonant-vowel] when singletons, and as [vowel-consonant-consonant-vowel] when geminates. The stress was placed on the first syllable. The vowels in the words were selected as the three Italian vowels [a, i, u].

The speech materials and measurements are described in section 1. Acoustic analyses carried out on the above speech materials are reported in section 2. In particular, time based parameters values are reported in section 2.1, and frequency and energy based parameters in section 2.2. In section 3, the results of the acoustic analyses are discussed. Section 3 also includes the conclusions and the indications for future work.

top of article

---

1. Speech materials and measurements

1.1 Data

In Italian, several disyllabic words form minimal pairs which can be distinguished on the sole basis of gemination of one consonant. For example, the words *casa*, 'home', and *cassa*, 'box', form a minimal pair words which are distinguished by the presence of either the singleton fricative [s] or the geminate fricative [s:]. Native speakers exhibit a natural attitude in producing disyllabic words of minimal pairs identified by the presence or absence of gemination of one consonant.

The above consideration led to the creation of a database formed by a set of vowel-consonant-vowel disyllabic words (the singleton case) and vowel-consonant-consonant-vowel disyllabic words (the geminate case) which would serve as the basis for studying gemination in all possible geminated consonantal forms of Italian. The words in the database included the entire set of those Italian consonants which appear in singleton and geminated forms in Italian *i.e.* [f, v, s, p, t, k, b, d, g, m, n, l, r]. The vowels in the words of the database were selected as the three cardinal vowels in Italian [a, i, u]. Consider that vowels in Italian form a larger set composed by [i, e, e, a, o, o, u].

Since it was expected from results of previous studies [2-5] that gemination would influence the values of acoustic parameters related to durational measurements, the words in the database were not included in a carrier phrase with the aim of avoiding the influence on parameters such as stress and intonation in a way which would be difficult to control. Six native speakers of Italian (three males and three females) uttered the speech materials described above. Each word was repeated three times. The speech materials considered in the present study belong to the above set of data with a restriction to fricatives which appear as both singleton and geminate in Italian and which are [f,: f, v,: v, s,: s].

The words analyzed were therefore 6 for each consonant (corresponding to the three vowels considered) and 6 for each speaker in three versions, leading to a total of 6 x 3 x 6 x 3 = 324 utterances (162 singleton utterances and 162 geminate utterances).

1.2 Subjects

Six Italian adult native speakers with no known articulatory impairment served as subjects in the experiment. They all spoke standard Italian. Some of them did not show any dialectal accent, while others presented the accent of the Italian region where they spent most of their life, *i.e*. Rome.

1.3 Recordings

The speech materials described above were produced by the speakers in a sound-treated room and recorded using a high quality recording system. The recordings were carried out at the Speech Laboratory of the INFOCOM Department at the University of Rome - La Sapienza. Each speaker produced the entire set of one repetition of words in a recording session. There were therefore three recording sessions for each speaker, corresponding to the three versions of the words. The subjects read the words to be pronounced from cards (presented to them by an operator) which were shuffled before each recording session. If a mistake occurred, the speaker was asked to repeat the word. The operator was a phonetically trained subject who also served as a controller on the quality of the produced speech sample. Therefore, if a word was judged to be unnatural it was asked to be repeated as well.

The set of words analyzed in the present study is reported in ***Table I.***

TABLE I The complete set of words analyzed

	F		V		S	
A	*afa*	*affa*	*ava*	*avva*	*asa*	*assa*
I	*ifi*	*iffi*	*ivi*	*ivvi*	*isi*	*issi*
U	*ufu*	*uffu*	*uvu*	*uvvu*	*usu*	*ussu*

The speech materials were then digitized using a software named UNICE by VECSYS which allows the use of appropriate oversampling factors in order to obtain a correct A/D conversion. The speech signals were filtered at 5 kHz, sampled at 10 kHz, and the samples were represented by 16 bits, before being stored in the memory of a PC.

1.4 Measurements

The UNICE software by VECSYS is a speech analysis program which accepts user commands to generate spectral displays of various types. The speech waveforms and their spectrograms can be presented on the screen of a PC. Spectrum slices at desired sampling points can also be obtained in standard forms such as the DFT (Discrete Fourier Transform) or the LPC (Linear Predictive Coding) spectra. In the analysis presented below, the DFT spectral analysis was selected, using a Hamming window of 256 samples corresponding to about 26 ms at a sampling rate of 10 kHz (the signal was pre-emphasized with a = 0.95).

1.4.1 Measurements in the time domain

The measurements carried out in the time domain corresponded in all cases to durational measurements. In particular the following parameters were selected (see ***Figure 1***):

1. Duration of the vowel preceding the consonant in the disyllabic word, indicated below as V1 duration. The vowel onset was identified by the appearance of a glottal pulse followed by other

regular glottal pulses. In those cases in which a glottal excitation was visible before regular vowel voicing, the vowel onset was taken as the beginning of regular vowel voicing, and the initial glottal excitation was discarded. Vowel offset was identified, by examination of both the waveform and the spectrogram, as the temporal sampling point where the glottal pulse disappeared and/or most frequency energy was in a frequency range over 1 kHz.

2. Duration of the vowel following the consonant in the disyllabic word, indicated below as V2 duration. The V2 onset was identified, by visual inspection of the spectrogram, as the temporal sampling point where the energy in the spectrogram appears in a frequency range above 1 kHz. The V2 offset was identified using the same procedure as V1 offset.
3. Duration of the consonant in the disyllabic word, indicated below as C duration. It should be noted that the analyzed consonants are characterized by the feature [+continuant], and therefore their duration can be considered as such, contrarily to what happens for stop consonants. Therefore, the consonant duration was determined as the interval between V1 offset and V2 onset.

1.4.2 Energy-based parameters and measurements in the frequency domain

Measurements in this context led to the definition of a set of energy-based parameters and a set of parameters corresponding to formants and their amplitudes in vowels, and f0 and its amplitude in vowels and consonants. In particular, the following energy-based parameters were considered:

1. $E_{totV1} = \sum_{i=t1}^{t2} |Xi|^2$ total energy of V1. X_i is the sample i, t_1 and t_2 are the temporal sampling points of vowel onset and vowel offset, respectively.
2. $P_{mV1} = \frac{E_{totV1}}{t2 - t1}$ average power of V1.
3. Total energy of C, indicated below as E_{totC} and calculated as for V1, with t_1 and t_2 that corresponding in the present case to V1 offset and V2 onset, respectively.
4. Average power of C, indicated below as P_{mC} and calculated as for the average power of V1.
5. Instantaneous energy at the CENTRE of V1, indicated as E_{iV1} and computed as for E_{totV1} but in a temporal window of 256 samples CENTREed in the middle of V1.
6. Instantaneous energy at the transition V1-C, indicated as $E_{iV1\text{-}C}$: The window of 256 samples is CENTREed on the temporal sampling point corresponding to V1 offset.
7. Instantaneous energy at the CENTRE of C, indicated as E_{iC} and calculated as E_{iV1}.
8. Instantaneous energy at C offset, indicated as E_{iCoff}: The last of the 256 samples corresponds to V2 onset samples.
9. Percentage of energy in the frequency band 0-350 Hz with respect of total energy sampled at the CENTRE of C.

All parameters listed above, except for the percentage at point 9, are in dB.

The formants (F1, F2, F3) and their amplitudes (A1, A2, A3), and f0 values and amplitude (A0) were estimated by visual inspection of DFT spectrum slices (obtained with UNICE) computed at different sampling points. In particular, the following parameters were extracted:

1. f0, A0, F1, A1, F2, A2, F3, A3, at the CENTRE of V1
2. f0, A0, F1, A1, F2, A2, F3, A3, at the offset of V1
3. f0, A0, F1, A1, F2, A2, F3, A3, at the transition from V1 to C
4. f0, A0, F1, A1, F2, A2, F3, A3, at the onset of V2
5. f0, A0, F1, A1, F2, A2, F3, A3, at the CENTRE of V2
6. f0 and A0 at the onset of voiced consonants
7. f0 and A0 at the CENTRE of voiced consonants
8. f0 and A0 at the offset of voiced consonants

Figure 1 shows, as an example, the different sampling points which were selected for the computation of frequency based parameters, and the corresponding computed parameters. As it can be noted, an extra frame in V1 to C transition (overlapping for one half with previous and for one half with following) was considered. Moreover, for each consonant, spectrum prints were carried out at onset, CENTRE and offset.

Spectrum prints and complete set of measurements described above can be found in [9].

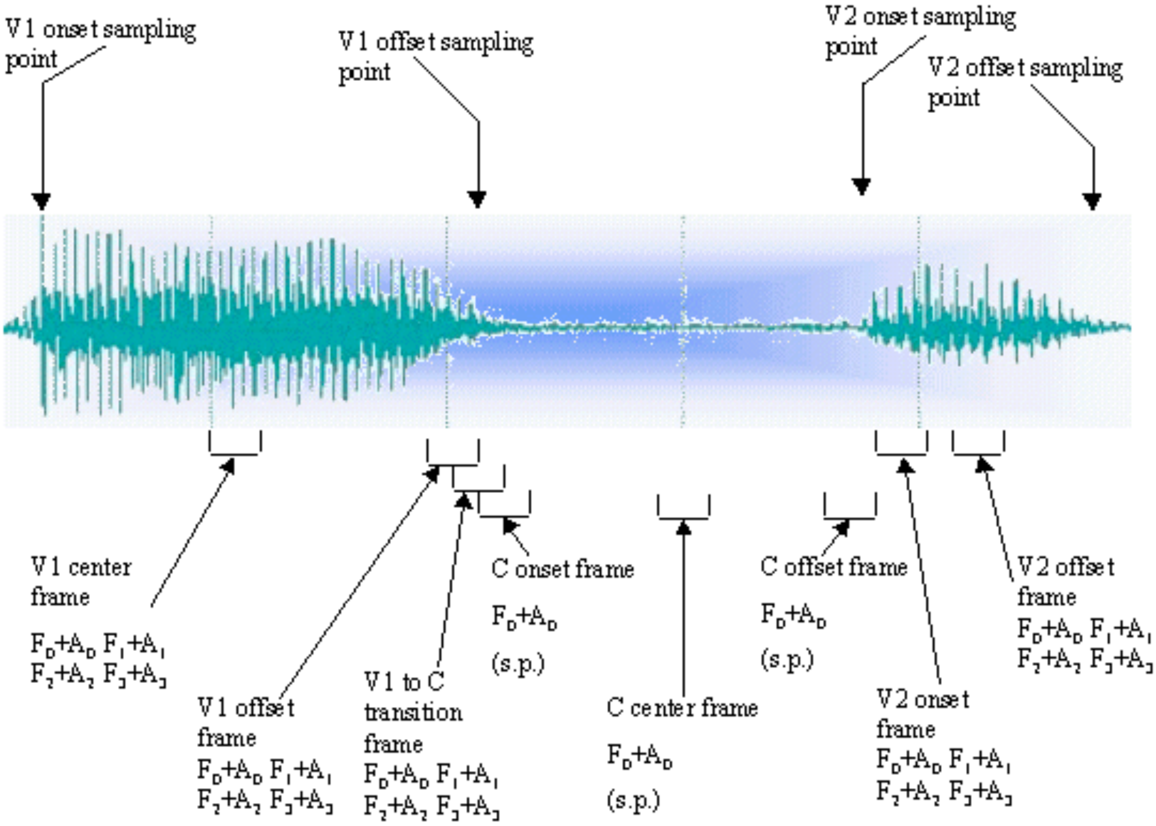

***Figure 1**. Sampling points which were selected for the computation of frequency-based*

parameters and the corresponding computed parameters (s.p. = spectrum printed).

[top of article](#)

2. Results of acoustic analyses

The aim of the acoustic analysis experiment was to understand whether the acoustic parameters described in 1.4.1 and 1.4.2 would show a significant difference in their value in singleton *vs*. geminate words. Results of the analyses are reported in the present paragraph. Statistical methods such as a Student ***t***-test, a maximum a-posteriori classification test, and the Spearman Rank Correlation Coefficient r_s were carried out in order to find out relationships among parameters, and to understand whether the values found were significantly different. Results of these tests are reported below.

2.1 Results in the time domain

The values of the acoustic time-domain parameters listed in section 1.4.1 were computed for each of the 162 singleton and 162 geminate utterances.

Results obtained are reported in ***Table II***, which contains the duration values of V1, C, and V2, as well as for the whole utterance, and in addition, the consonant *vs*. preceding vowel ratio, averaged over all repetitions and speakers, for each consonant in the three vowel contexts [a, i, u] in geminate and singleton forms. ***Table II*** also includes the standard deviation values obtained in correspondance to the above average values.

TABLE II Durational measurements (in ms) of all the analyzed utterances. Average values (and standard deviations) with respect of all repetitions and speakers are reported. Average duration of V1 (V1dm), average duration of C (Cdm), average duration of V2 (V2dm), average duration of complete utterance (UTdm), average Cd/V1d (Rm).

	$V1d_m$	(StD)	Cd_m	(StD)	$V2d_m$	(StD)	UTd_m	(StD)	Rm	(StD)
AFA	165.8	*18.8*	151.7	*21.4*	111.6	*28.6*	429.0	*37.0*	0.93	*0.2*
AFFA	123.2	*18.1*	248.3	*30.3*	109.1	*22.4*	480.7	*45.6*	2.06	*0.4*
AVA	188.7	*21.4*	83.3	*13.5*	123.0	*26.0*	395.0	*45.0*	0.45	*0.1*
AVVA	126.5	*20.7*	205.8	*27.0*	108.0	*16.2*	440.3	*36.6*	1.69	*0.4*
ASA	175.9	*17.0*	147.2	*13.9*	122.6	*27.3*	445.8	*37.1*	0.85	*0.2*
ASSA	125.3	*20.4*	250.1	*35.6*	113.7	*24.3*	489.2	*41.0*	2.08	*0.6*
IFI	164.3	*23.1*	153.0	*30.7*	109.9	*22.9*	427.3	*36.4*	0.96	*0.3*
IFFI	115.1	*27.8*	253.5	*37.4*	112.1	*27.0*	480.7	*49.2*	2.38	*0.9*

IVI	185.0	*26.6*	90.3	*13.1*	118.8	*23.6*	394.1	*48.7*	0.50	*0.1*
IVVI	122.4	*28.8*	202.1	*29.6*	117.5	*29.3*	442.0	*61.5*	1.75	*0.5*
ISI	175.5	*21.6*	164.2	*29.2*	115.1	*23.8*	454.8	*34.6*	0.97	*0.3*
ISSI	124.7	*27.3*	260.0	*36.1*	113.6	*20.0*	498.3	*44.9*	2.19	*0.6*
UFU	163.8	*34.7*	163.7	*26.2*	118.4	*21.6*	446.0	*43.1*	1.09	*0.5*
UFFU	120.3	*28.8*	253.2	*38.1*	109.8	*18.4*	483.2	*37.1*	2.27	*0.8*
UVU	188.2	*34.0*	105.4	*19.1*	135.2	*23.4*	428.8	*49.1*	0.58	*0.2*
UVVU	156.4	*29.0*	171.3	*34.3*	134.7	*27.2*	462.4	*48.1*	1.15	*0.4*
USU	173.8	*18.7*	155.3	*25.8*	115.4	*25.5*	444.5	*40.7*	0.91	*0.2*
USSU	125.2	*25.2*	255.0	*39.6*	108.6	*24.2*	488.8	*39.5*	2.15	*0.7*

As it can be noted from the data reported in ***Table II***, there are two parameters which have quite different values in singleton *vs*. geminate words; these are V1 duration and C duration. On the other hand, V2 duration in singleton *vs*. geminate forms does not vary as much. The same observation can be made for the duration of the whole utterance. In particular, note that V1 duration was always higher in singletons than in geminates while the opposite effect was found for C.

The significance of the differences observed between the average values of V1 and C duration in singleton *vs*. geminate words was tested by application of a Student t-test for independent groups. In particular, we considered the ***t***-statistic $t = \frac{\overline{x_1} - \overline{x_2}}{S_{\overline{x_1} - \overline{x_2}}}$ with n1 = n2 = 6 degrees of freedom corresponding to the number of speakers, and thus without considering as independent the three repetitions of each speaker (this hypothesis corresponds to the worst case) [8]. Results of this test are reported in ***Table III***.

TABLE III Result of the Student ***t***-test. For each vowel and consonant the ***t***-value (observed) was computed. The null hypothesis can be rejected for the averages of the two groups (singleton and geminate) at the corresponding *p* level of significance.

		A		**I**		**U**	
		V1	**C**	**V1**	**C**	**V1**	**C**
F	***t*** value	4,00	6,38	3,33	5,09	2,36	4,74
	p value	2,5E-03	8,0E-05	7,6E-03	4,7E-04	4,0E-02	7,9E-04
		V1	**C**	**V1**	**C**	**V1**	**C**
V	***t*** value	5,12	9,94	3,91	8,46	1,74	4,11
	p value	4,5E-04	1,7E-06	2,9E-03	7,2E-06	1,1E-01	2,1E-03
		V1	**C**	**V1**	**C**	**V1**	**C**
S	t value	4,67	6,60	3,57	5,05	3,79	5,17
	p value	8,8E-04	6,1E-05	5,1E-03	5,0E-04	3,5E-03	4,2E-04

It is clear that the result is strongly significant: for V1 one finds always $p < 0.01$ except in the case UFU/UFFU ($p < 0.05$) and UVU/UVVU ($p < 0.12$), while for C the value was always $p < 0.001$ except, again, in the case UVU/UVVU ($p < 0.003$).

An interesting result is presented in ***Table IV***, which shows the correlation coefficient (r_s) for the singleton group and the geminate group, first separately and then combined. In section 4 we will return to comment this table.

TABLE IV Spearman Rank Correlation Coefficient r_s calculated among the time related parameters; once on the two groups of singleton and geminate separately, twice on all utterances.

	V1 sin	C sin	V2 sin	V1 gem	C gem	V2 gem
V1 sin	1.00	-0.38	0.53	not significant		
C sin	-0.38	1.00	-0.26			
V2 sin	0.53	-0.26	1.00			
V1 gem	not significant			1.00	-0.46	0.65
C gem				-0.46	1.00	-0.26
V2 gem				0.65	-0.26	1.00

	V1	C	V2
V1	1.00	-0.78	0.55
C	-0.78	1.00	-0.33
V2	0.55	-0.33	1.00

Finally, ***Table V*** shows the results of the application of the Maximum Likelihood Criterion [8]: one dimensional applied on the time related parameters Cd and Cd/V1d; two dimensional on the plane (Cd, V1d). It can be observed that the classification in one dimension on the C/V1 ratio leads to very good results, when compared to the other two parameters considered. In particular one can note that:

- if the ratio is used, almost all errors are due to one or two speakers, while with the other parameters a larger dispersion is found;
- it was found, by a trial-and-error procedure that the minimum number of errors (7%) is obtained with a ratio equal to 1.05, which is different from the equiprobability point given by MLC which is equal to 1.3.

These results will be discussed in section 3. In addition, a comparison with stops will be carried out.

TABLE V Results of the uni-bi-dimensional Maximum Likelihood Criterion used for classifying singleton *vs.* geminate. E.P.P. = Equal Probability Point

	UNI MLC						BI MLC	
	on Cd			on Cd/V1d			on Cd and V1d	
Context	E.P.P.	Errors	Error %	E.P.P.	Errors	Error %	Errors	Error %
Overall	182	39/324	12.0	1.30	39/324	12.0	34/324	10.5
Male	176	12/162	7.4	1.13	5/162	3.0	6/162	3.7
Female	188	24/162	14.8	1.44	35/162	21.6	35/162	16.6
[a]	180	6/108	5.5	1.21	9/108	8.3	9/108	8.3
[i]	186	13/108	12.0	1.34	11/108	10.2	11/108	9.3

[u]	183	20/108	18.5	1.37	21/108	19.4	21/108	15.7
[f]	199	6/108	5.5	1.54	13/108	12.0	13/108	6.5
[v]	131	6/108	5.5	0.78	7/108	6.5	7/108	5.5
[s]	198	10/108	9.3	1.34	9/108	8.3	9/108	7.4

2.2 Results in the frequency domain

No significant effect was found on the parameters in the frequency domain. The only parameters which seem to be influenced by the presence of gemination are the fundamental frequency (f0) and, in a less important manner, the first formant (F1). As regards f0, its variation is equal to about 10 Hz in the geminate form (15%) in those points where V1 was sampled, and of about 4 Hz in V2. These variations are small, considering that the standard deviation is greater than 40 Hz (>25%). The only valid observation which can be carried out is that this result confirms the temporal-parameters analysis result, *i.e.* that V1 appears as more influenced by the presence of gemination with respect to V2.

All the other parameters do not seem to be related to gemination.

For the sake of completeness, ***Tables VI*** and ***VII*** report the data on the frequency based parameters (energy, formants, and f0) averaged over all utterances. Details on the above parameters can be found in [9]. Since they were not significant, they were omitted in the present paper in order to limit the length of the manuscript.

TABLE VI Energy-based parameters. Mean values and Standard Deviation with respect of all the repetitions, speakers, vowels and consonants.

	EtotV1	**PmV1**	**EtotC**	**PmC**	**EiV1cent.**	**EiV1-C**	**EiCcent.**	**EiCoffset**	**CCENTRE% 0-350/0-5000**
Singleton	94.62	62.17	78.64	47.52	86.96	78.68	69.40	69.27	10.44
(StD)	*5.494*	*5.322*	*4.398*	*5.104*	*5.845*	*4.867*	*5.812*	*5.737*	*21.564*
Geminate	94.12	63.19	79.69	46.09	88.49	79.02	68.04	68.27	12.22
(StD)	*5.871*	*5.502*	*4.687*	*5.015*	*5.771*	*4.294*	*6.414*	*5.700*	*23.656*

TABLE VII Frequency-based parameters: Average and Standard Deviation with respect of all repetitions, speakers, vowels and consonants.

	V1 CENTRE							
	f0	**A0**	**F1**	**A1**	**F2**	**A2**	**F3**	**A3**
Singleton	158.93	12.21	543.04	32.93	1555.28	29.79	2909.18	24.72
(StD)	*44.724*	*6.409*	*301.698*	*8.847*	*791.763*	*8.680*	*497.112*	*11.664*
Geminate	164.25	13.07	550.03	34.58	1544.82	31.40	2859.14	24.46

(StD)	*45.201*	*7.212*	*298.433*	*8.961*	*762.853*	*8.288*	*453.889*	*10.265*
	V1 OFFSET							
	f0	**A0**	**F1**	**A1**	**F2**	**A2**	**F3**	**A3**
Singleton	146.05	11.53	493.97	28.27	1547.35	24.05	2763.89	19.42
(StD)	*43.262*	*5.430*	*250.944*	*7.717*	*737.406*	*7.002*	*526.358*	*7.487*
Geminate	156.42	13.36	516.21	28.91	1547.79	26.44	2796.79	20.94
(StD)	*45.614*	*9.644*	*253.632*	*8.016*	*741.715*	*7.824*	*474.517*	*8.235*
	V1 TO C TRANSITION							
	f0	**A0**	**F1**	**A1**	**F2**	**A2**	**F3**	**A3**
Singleton	141.98	10.71	463.07	23.09	1539.42	20.82	2712.70	17.20
(StD)	*42.627*	*4.998*	*220.652*	*7.248*	*713.576*	*6.422*	*517.170*	*6.652*
Geminate	151.11	11.42	487.89	23.63	1542.70	21.20	2779.52	18.19
(StD)	*45.326*	*4.925*	*228.432*	*7.225*	*696.695*	*6.910*	*434.672*	*6.658*
	C ONSET - CENTRE - OFFSET							
	f0	**A0**		f0	**A0**		f0	**A0**
Singleton	134.17	9.89		126.13	9.06		126.02	8.50
(StD)	*36.010*	*5.053*		*34.070*	*4.426*		*29.676*	*4.853*
Geminate	143.15	9.83		125.37	10.81		125.18	8.92
(StD)	*39.297*	*4.649*		*31.036*	*16.200*		*28.846*	*4.827*
	V2 ONSET							
	f0	**A0**	**F1**	**A1**	**F2**	**A2**	**F3**	**A3**
Singleton	133.74	8.05	444.38	26.23	1515.11	23.12	2763.73	18.60
(StD)	*29.302*	*4.221*	*192.772*	*7.114*	*704.201*	*7.422*	*425.909*	*7.159*
Geminate	136.03	8.64	434.82	26.59	1521.07	23.98	2749.64	19.13
(StD)	*30.380*	*4.375*	*166.065*	*7.017*	*667.207*	*6.891*	*414.415*	*7.446*
	V2 CENTRE							
	f0	**A0**	**F1**	**A1**	**F2**	**A2**	**F3**	**A3**
Singleton	128.73	9.44	492.92	29.68	1524.60	24.91	2846.87	19.24
(StD)	*29.898*	*5.423*	*262.296*	*6.848*	*755.361*	*7.771*	*479.562*	*7.742*
Geminate	132.33	9.39	490.87	31.26	1533.18	26.76	2771.12	19.83
(StD)	*31.357*	*5.164*	*256.818*	*6.796*	*733.358*	*7.351*	*435.631*	*8.491*

top of article

---

3. Discussion of acoustic analyses results

3.1 Effect of gemination

Results of acoustic analyses showed that, as reported in section 3, frequency-related parameters were not significantly different in singletons *vs*. geminates. This observation was made on the basis of measurements of formants and their amplitudes, f0 and its amplitude, sampled at different sampling times within the word. These spectral parameters were estimated within the vowels and consonants, as well as at the transitions from vowel to consonant and *vice versa*. A similar observation was made in the case of energy related parameters. The energy and power values computed in different ways, *i.e*. instantaneous and averaged, did not show any evidence for a correlation with the singleton-geminate distinction.

Contrarily, some of the time related parameters showed strong evidence for a correlation with gemination. In particular, two parameters, V1 duration and C duration, were significantly different in singletons *vs*. geminates. The effect observed was of an elongation of the geminate consonant with respect of the singleton consonant, and a shortening of the vowel preceding the consonant in geminate *vs*. singleton words.

By observing the data reported in ***Table IV*** of the preceding section, some interesting conclusions can be made. Indeed, it is clear that, within the group of singletons and within the group of geminates, *i.e*. by keeping the two classes separate (matrix on the left of the ***Table***) a significant correlation between the durations of the segments is present. In particular, an increase in the duration of the consonant leads to a shortening of the two vowels V1 and V2, and *vice versa*, which means that this effect is present also in absence of gemination.

When all the utterances are considered in one group (matrix on the right of ***Table IV***), one observes that the correlation between C and V1 duration is about -0.78 (double the value found with separated classes) while the correlation between C and V2, already present in the first case, remains almost unchanged in the second case.

The above observations support the use and give credit to analyses of relative time parameters, in the search for possible invariants.

3.2 Comparison of acoustic correlates of gemination in fricatives *vs.* stops

The average difference in V1 duration was 49 ms (» -28%) and in C duration 98 ms (» +73%). This result is in agreement with previous studies on gemination in Italian, which focused mostly on the effect of gemination in stop consonants. In particular, it was found that the average difference in stops [5] was 43 ms for V1 duration (» -26%) and 92 ms for the stop closure duration (+101%).

TABLE VIII Mean values and Standard Deviation of the time related parameters with respect of all the repetitions, speakers, consonants and vowels.

		V1d	Cd	V2d	UTd	Cd/V1d	

	Singleton	175.66	134.91	118.90	429.46	0.80	
	(StD)	*25.871*	*37.602*	*25.248*	*45.567*	*0.333*	
	Geminate	126.58	233.25	114.12	473.96	1.97	
	(StD)	*27.145*	*45.066*	*24.288*	*48.505*	*0.699*	

By building up a durational correlation coefficient matrix as in ***Table IV*** but referred to the stops, it was found that similar results appear.

The MLC application shows the Equal Probability Points (E.P.P):

1. in fricatives (see Table V) the E.P.P is equal to 182 ms (%errors=12%) for Cd, and E.P.P is equal to 1.30 (%errors=12%) for the Cd/V1d ratio;

2. in stops [5] E.P.P is equal to 128 ms (%errors=4%) for CLd and 0.93 (%errors=8%) for CLd/V1d ratio.

The above difference, already pointed out by Bertinetto and Vivalda [7] finds a first justification in the characteristics of stops which are [-continuant] with respect to fricatives which are [+continuant]. This property causes a fricative to be on average longer than a stop.

Finally, it was found that the durations Cd and CLd were too much different to allow a classification on this parameter, while the opposite is true for the ratio C/V. In fact, an a-posteriori classification based on this last parameter with a boundary set at r = C/V = 1.02 leads to a surprising result of 8% of errors on stops (as with the MLC) and 8% of errors with fricatives (*i.e.* better than with the MLC!).

As a general conclusion, the two most relevant outcomes of the present work can be summarized as follows.

- First, the general tendency of shortening the pre-consonant vowel and of lengthening the consonant in a geminate utterance which was observed on stops in previous studies, was confirmed for fricatives. Moreover, the ratio of these two values appeared as significant and as a valid way of synthesizing the two effects in only one parameter. However, a more careful examination of the speech materials under study, highlighted that a degree of correlation between the two afore mentioned durational parameters is inherently present in singleton utterances on the one hand and geminates on the other hand; that is to say that the effects observed in singletons *vs*. geminates are already present in singletons *vs*. singletons and in geminates *vs*. geminates. As an additional comment on this point, the same effect was observed for durational parameters of the other segments in the utterance, as for example the post-consonantal vowel duration.

  Whereas the effect appears as a general tendency related to the particular language under examination, it is important to point out that it is reinforced in geminates *vs*. singletons, and that the degree of correlation in this case is higher. The above result is important since it quantifies a hypothesis suggested in a previous work on the gemination of consonants [4] that the effect observed can be justified by the need of preserving the rhythmical structure of a word including a geminate. An additional confirmation to this hypothesis would be possible by integrating the word in a sentence uttered at different speaking rates. This will be the first object of future work.

- Secondly, whereas the ratio between the duration of the consonant and pre-consonant vowel was slightly different in the case of fricatives with respect to stops, a single value was found which was capable of discriminating singletons *vs*. geminates, for both stops and fricatives, in a very satisfactory way. The above value, leading to an 8 % of correct classification score for both fricatives and stops, was equal to about 1. This result suggests that speaker intention when pronouncing a geminate is to produce a consonant which is at least as long as the vowel preceding it. The second object of future work will be to further investigate the above interpretation.

top of article

References

[1] Muljacic Z. **(1972)**. *Fonologia della lingua Italiana*. (Il Mulino: Bologna)

[2] Rossetti R. **(1994)** Gemination of Italian stops, *Journal of the Acoustical Society of America* 95,2 SP25, p.2874

[3] Rossetti R. **(1993)**. C*aratteristiche acustiche del fenomeno di geminazione nelle consonanti occlusive italiane: applicazione all'adattamento automatico di pronunce straniere*, Degree Thesis in Electronics Engineering, Rome University, La Sapienza

[4] Shrotriya N., Siva Sarma A.S., Verma R., Agrawal S.S. **(1995)**. Acoustic and perceptual characteristics of geminate Hindi stop consonants, in *Proceedings of ICPhS95*, edited by K.Elenius and P.Branderud, 4. Arne Strombergs Grafiska: Stockolm, pp.132-135

[5] Esposito A., Di Benedetto M.G. **(1997)**. Acoustic and Perceptual Study of Gemination in Italian Stops, under review in *Journal of the Acoustical Society of America*

[6] Blumstein S.E., Pickett E. Burton M. **(1998)**. *Effects of speaking rate on Singleton/Geminate consonant contrast in Italian*, unpublished manuscript

[7] Bertinetto P.M., Vivalda E. **(1978)**. Recherches sur la perception des oppositions de quantité en italien, *Journal of Italian Linguistics* 3, pp.97-116

[8] Dillon W.R., Goldstein M. **(1984)**. *Multivariate Analysis*. Wiley J. & Sons

[9] Giovanardi M. **(1998)**. *Analisi Acustica e Sintesi delle consonanti fricative singole e geminate in Italiano*, Degree Thesis in Electronics Engineering, Rome University, La Sapienza

---

---

top of article

***web-sls* front page**

web-sls *article:* #98.01
published: 20th July 1998
archived: n/a

academic comments to: the author
publishing comments to: Mark Tatham, Technical and Publishing Editor