Acoustic Analysis of Singleton and Geminate Affricates in Italian

by

Simone Faluschi and Maria-Gabriella Di Benedetto

INFOCOM Department, Univresity of Rome 'La Sapienza'

via Eudossiana, 18, 00184 Roma
simone.fal@tiscalinet.it
gaby@acts.ing.uniroma1.it

Abstract

The aim of this study was to analyze gemination of affricate consonants in Italian. Disyllabic VCV (vowel-consonant-vowel) and VCCV (vowel-consonant-consonant-vowel) words were considered, where C was one of the four affricate Italian consonants [tʃ, dʒ, ts, dz], in singleton and geminate forms, while V was one of the three Italian point vowels [a, i, u]. Time-related, frequency-related, and energy-related parameters were estimated; These parameters were measured at different sampling times, in order to observe their variation within an utterance. Results showed that the preconsonant vowel and consonant durations varied significantly among singleton vs. geminate utterances. Frequency and energy parameters did not present differences as large as durational parameters, although they were significantly different in the two groups. While results in the time-domain are in agreement with previous studies on Italian gemination of consonants (stops, liquids, fricatives and nasals), frequency and energy significant variations among single vs. geminate groups seem to be peculiar to the class of affricates.

Introduction

In Italian, minimal pairs, formed by words which can be identified only by consonant gemination, are common. Words belonging to minimal pairs are orthographically distinguished by a double grapheme of the geminate consonant (for example: micia (pussy-cat) and miccia (fuse)). A problem, which is still unsolved, regards the identification of acoustic correlates of singleton vs. geminates, and their perceptual verification. In order to give an answer to this question in Italian, several analyses - within the GEMMA project in progress at the INFOCOM Department at University 'La Sapienza' in Rome - were performed on the following classes of consonants: stops, liquids, fricatives and nasals. Results on stop consonants were reported in Rossetti (1993, 1994) and Esposito and Di Benedetto (1999). Liquids were analyzed in Argiolas et al. (1995), while the analysis of fricatives was reported in Giovanardi (1998) and Giovanardi and Di Benedetto (1998). Finally, results on nasal consonants were reported in Mattei (1999) and Mattei and Di Benedetto (2000).

As already pointed out in (Mattei and Di Benedetto, 2000) gemination has been analyzed in other languages as well: (Shrotriya, 1995), (Blumstein et al., 1998), (Rochet and Rochet, 1995), (Cohn et al., 1999), (Abramson, 1999) (Local and Simpson, 1999), (Arvaniti, 1999), (Louali and Maddieson, 1999). However, the above studies do not analyze affricate consonants. The only study referring to this particular class of consonants is (Abramson, 1999); However, the author points out that a pre-test perceptive analysis presented a percentage of mistakes which was too high, and therefore, the affricates were discarded.

The paper is organized as follows: 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 Section 3, the results of the acoustic analyses are discussed. Section 3 also includes the conclusions and the indications for future work.

SECTION 1

Speech materials and measurements

In Italian, several disyllabic words form minimal pairs which can be distinguished on the sole basis of gemination of one consonant. Native speakers exhibit a natural attitude in producing disyllabic words of minimal pairs identified by the presence or absence of consonant gemination. The above consideration led to the creation of the GEMMA 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. Affricate consonants were studied when coarticulated with the three Italian point vowels [a, i, u] and represent a subset of

Italian vowels [a, e, ɛ, i, o, ɔ, u]. Words were pronounced by six pronunciation defectless and dialectal inflexionless native speakers, three men and three women aged between twenty-four and fifty. In order to keep total control on the parameters used, words were pronounced in isolation and not in carrier sentences. By this way, elements such as intonation or internal stress, which could have had an influence, were eliminated. Each word was pronounced three times by each speaker. Our procedure in building up the database was motivated by the strong need of having data with well-controlled parameters, as a starting point for future more realistic settings. The analyzed words

in the present study were therefore 3 for each affricate consonant (which are [tʃ, dʒ, ts, dz] and their geminate version) and 6 for each speaker in three repetitions, leading to a total of 3x4x2x6x3=432 utterances (216 singletons and 216 geminates). Words were written on cards which were presented to the speaker by the operator. Cards were shuffled after each recording session.

All the utterances were produced and recorded at the Speech Laboratory of the INFOCOM Department at the University of Rome 'La Sapienza' (Italy). Equipment used for this experiment is of a superior quality and the recording took place in a sound-treated room. The operator who supervised the recording procedure was an acoustically trained subject. When there was an evident pronunciation mistake or if the pronunciation was judged unnatural, the speaker was compelled to repeat the word. There were three recording sessions, corresponding to each of the three repetitions.

The set of analysed words is reported in Table I. Geminate utterances are indicated by a double grapheme of the consonant.

	tʃ		dʒ		ts		dz	
a	atʃa	attʃa	adʒa	addʒa	atsa	attsa	adza	addza
i	itʃi	ittʃi	idʒi	iddʒi	itsi	ittsi	idzi	iddzi
u	utʃu	uttʃu	udʒu	uddʒu	utsu	uttsu	udzu	uddzu

TABLE I The complete set of analyzed words

After recording, the utterances were digitized using the UNICE software produced by VECSYS (Vecsys, 1989). Speech signals were filtered at 5 kHz, sampled at 10 kHz, and each sample was represented by using 16 bits. The resulting signals were stored on a PC. Next, an analysis performed by UNICE included spectrograms, DFT (Discrete Fourier Transform) and LPC (Linear Predictive Coding) spectra. For this processing, a pre-emphasizing filter (with $\alpha=0.95$) on the signal and a Hamming window of 256 samples were used.

The GEMMA project considers the measurement of standard parameters (Giovanardi and Di Benedetto, 1998) and (Mattei and Di Benedetto, 1999). Since the affricate consonants are characterized by having an occlusive and a fricative phase, it was decided to enlarge the set of analyzed parameters. The following parameters were used in the present study:

1. duration of the pre-consonant vowel, indicated as V1d. 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

2. duration of the occlusive part of consonant, indicated as C1d

3. consonant vowel, indicated as V1d. 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 examining both the waveform and the spectrogram

4. duration of the occlusive part of consonant, indicated as C1d

5. duration of the fricative part of consonant, indicated as C2d

6. duration of the consonant, indicated as Cd (= C1d + C2d)

7. duration of the vowel following the consonant, indicated as V2d. The V2 onset was identified, by visual inspection of both the waveform and the spectrogram. The V2 offset was identified as the temporal sampling point where the glottal pulse disappeared

8. duration of the whole utterance, indicated as Utd

9. 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

$$E_{totV1} = \sum_{i=t1}^{t2} |Xi|^2$$

10. average power of V1

$$P_{mV1} = \frac{E_{totV1}}{t2 - t1}$$

11. total energy of occlusive phase of C, indicated as E_{totC1} and computed as for V1, with t_1 and t_2 corresponding to V1 offset (C1 onset) and C2 onset (C1 offset), respectively

12. average power of occlusive phase of C, indicated as P_{mC1} and computed as for the average power of V1

13. total energy of fricative phase of C, indicated as E_{totC2} and computed as for V1, with t_1 and t_2 that corresponding to C1 offset (C2 onset) and C2 offset, respectively

14. average power of fricative phase of C, indicated as P_{mC2} and computed as for the average power of V1

15. total energy of the whole consonant, indicated as E_{totC} and computed as described above

16. average power of the whole consonant, indicated as P_{mC}

17. instantaneous energy at V1 center, indicated as E_{iV1}, computed in a window of 256 samples centered on V1

18. instantaneous energy at V1-C1 transition, indicated as E_{iV1-C1}: The window of 256 samples is centered on V1 offset

19.   Instantaneous energy at C1 center, indicated as E_{iC1} and computed as E_{iV1}

20.   instantaneous energy at C1-C2 transition, indicated as E_{iC1-C2}: The window of 256 samples is centered on C1 offset (C2 onset)

21.   instantaneous energy at C2 center, indicated as E_{iC2} and computed as E_{iV1}

22.   instantaneous energy at C2 offset, indicated as $E_{iC2off,}$ computed right before the first 256 samples of V2 onset

23.   F0, A0, F1, A1, F2, A2, F3, A3, at V1 center (where F1,F2 and F3 are the formants and A1, A2 and A3 their amplitudes)

24.   F0, A0, F1, A1, F2, A2, F3, A3, at V1 offset

25.   F0, A0, F1, A1, F2, A2, F3, A3, at the transition from V1 to C

26.   F0 and A0 at the onset of voiced consonants

27.   F0 and A0 at the centre of occlusive phase of voiced consonants

28.   F0 and A0 at the centre of fricative phase of voiced consonants

29.   F0 and A0 at the offset of voiced consonants

30.   F0, A0, F1, A1, F2, A2, F3, A3, at V2 onset

31.   F0, A0, F1, A1, F2, A2, F3, A3, at V2 center

Sampling points selected for the computation of the acoustic parameters and indication of the frames where frequency parameters were estimated are in Figure 1 (*MISSING FIGURE*).

Time and energy domain parameters were not normalized since the comparisons were between values obtained by the same speakers who pronounced the singleton and geminate utterances under the same conditions. Moreover, normalization would have obscured potentially relevant energy information.

SECTION 2

2 Results of acoustic analysis

In this paragraph, time, frequency and energy related parameters are reported. The following statistical tests were performed: Mono and Multivariate Anova, the maximum a-posteriori classification test, the maximum likelihood classification test and the Spearman Rank Correlation Coefficient **r_s**. Results in time and frequency domains are reported separately.

2.1 Results in the time domain

Table II reports the average values of V1d, C1d, C2d, V2d and Utd for each of the four consonants coarticulated with the three vowels, and the corresponding standard deviations. As can be noticed on Table II, parameters which present relevant differences between singleton and geminate utterances are: V1d, C1d and C2d (as a reminder, C1 and C2 together form the affricate). A general tendency to shorten the first vowel duration and to lenghten the consonant duration (both in occlusive and fricative phase) is present. On the contrary, there are no relevant variations in the second vowel duration. Statistical analyses on the data confirm the above observations and are reported in Table III. The values which significantly vary (p<0.05) between singletons to geminates, are indicated in bold characters. A strong significant variation is found for V1d, C1d and C2d, except for [ʦ] coarticulated with [i]. The parameter Utd increases significantly in rather all geminate utterances. Moreover, the second vowel duration V2d does not vary significantly between singletons and geminates, except for [ʤ] articulated with [a].

	V1d	C1d	C2d	V2d	Utd		V1d	C1d	C2d	V2d	Utd		V1d	C1d	C2d	V2d	Utd
atʃa	160.0	73.1	100.9	112.3	446.3	itʃi	137.4	64.0	122.4	104.6	428.4	utʃu	163.6	66.0	103.7	131.7	465.0
(StD)	27.6	34.7	20.5	19.6	43.8	(StD)	20.8	29.2	16.2	17.9	29.8	(StD)	27.4	37.9	24.0	23.7	32.0
attʃa	113.2	137.8	128.7	107.5	487.2	ittʃi	99.3	122.8	158.4	110.7	491.3	uttʃu	110.9	151.1	123.0	125.0	509.9
(StD)	19.2	13.9	28.1	12.2	29.3	(StD)	17.9	20.4	26.1	21.0	37.5	(StD)	25.4	39.4	24.7	22.4	51.7
adʒa	169.0	92.0	49.1	142.3	452.3	idʒi	166.7	95.9	52.6	141.6	456.8	udʒu	173.5	85.7	44.1	146.1	449.5
(StD)	20.6	18.9	13.6	26.1	47.4	(StD)	28.3	17.5	15.7	30.6	53.4	(StD)	32.1	21.1	16.5	26.5	45.0
addʒa	127.3	156.1	61.5	125.9	470.9	iddʒi	111.7	162.1	74.1	129.4	477.3	uddʒu	120.2	154.0	61.3	137.3	472.8
(StD)	16.0	17.7	11.0	15.9	42.2	(StD)	21.3	28.2	25.5	30.6	56.6	(StD)	21.6	21.3	20.8	29.9	67.7
atsa	121.3	89.6	129.8	109.9	450.6	itsi	106.7	84.4	149.6	109.7	450.4	utsu	133.2	73.3	140.7	115.3	462.5
(StD)	23.3	11.0	34.0	23.1	37.0	(StD)	25.9	20.2	31.3	18.1	32.2	(StD)	30.6	26.9	22.4	16.3	41.1
attsa	106.0	112.2	167.0	117.4	502.6	ittsi	94.5	114.0	171.0	123.2	502.7	uttsu	103.8	96.3	178.8	115.1	493.9
(StD)	18.7	18.8	22.0	20.6	43.5	(StD)	17.9	31.4	34.7	22.8	48.0	(StD)	21.9	20.4	19.4	15.8	40.4
adza	163.4	89.9	78.6	139.7	471.7	idzi	148.4	85.9	90.9	148.1	473.4	udzu	150.8	81.6	80.9	139.7	453.0
(StD)	24.7	13.5	19.3	18.9	42.9	(StD)	37.5	16.5	21.6	20.7	35.7	(StD)	23.7	18.8	18.1	23.8	44.8
addza	127.8	139.8	102.3	136.3	506.2	iddzi	104.7	136.5	120.2	139.7	501.1	uddzu	117.7	116.8	112.3	136.4	483.1
(StD)	24.5	35.3	23.0	29.0	57.4	(StD)	23.9	36.4	38.1	19.0	53.0	(StD)	17.1	26.9	29.4	20.0	43.1

TABLE II Average values (and standard deviations) of V1d, C1d, C2d, V2d and Utd, over all repetitions and speakers. All values are in msecs.

		A					I					U				
		V1d	C1d	C2d	V2d	Utd	V1d	C1d	C2d	V2d	Utd	V1d	C1d	C2d	V2d	Utd
tʃ	F ratio	**34.89**	**53.90**	**53.90**	0.78	**10.83**	**34.69**	**49.22**	**24.66**	0.88	**31.05**	**35.75**	**43.56**	**5.62**	0.78	**9.82**
	p value	**0.0000**	**0.0000**	**0.0018**	0.3842	**0.0023**	**0.0000**	**0.0000**	**0.0000**	0.3554	**0.0000**	**0.0000**	**0.0000**	**0.0235**	0.3845	**0.0036**
dʒ	F ratio	**46.04**	**110.50**	**9.15**	**5.14**	1.54	**43.36**	**71.55**	**9.23**	1.43	1.25	**34.23**	**93.40**	**7.55**	0.88	1.48
	p value	**0.0000**	**0.0000**	**0.0047**	**0.0298**	0.2231	**0.0000**	**0.0000**	**0.0046**	0.2394	0.2722	**0.0000**	**0.0000**	**0.0096**	0.3546	0.2328
ts	F ratio	**4.73**	**19.22**	**15.15**	1.06	**14.91**	2.69	**11.29**	3.80	3.86	**14.75**	**10.96**	**8.32**	**29.77**	0.00	**5.36**
	p value	**0.0367**	**0.0001**	**0.0004**	0.3094	**0.0005**	0.1102	**0.0019**	0.0595	0.0576	**0.0005**	**0.0022**	**0.0068**	**0.0000**	0.9679	**0.0268**
dz	F ratio	**18.84**	**31.39**	**11.15**	0.17	**4.19**	**17.39**	**28.90**	**8.04**	1.62	3.39	**23.24**	**20.71**	**14.83**	0.20	**4.22**
	p value	**0.0001**	**0.0000**	**0.0020**	0.6785	**0.0485**	**0.0002**	**0.0000**	**0.0077**	0.2120	0.0745	**0.0000**	**0.0001**	**0.0005**	0.6587	**0.0477**

TABLE III Results of the Anova test performed on V1d, C1d, C2d, V2d and Utd. F-values are listed. The null hypothesis can be rejected at the p level of significance indicated on the table. Bold characters indicate significantly different values.

A Spearman Rank correlation test was carried out in order to verify if the relation between phonemes durations, in particular those between V1d, C1d and C2d, was based on gemination. The results were as follows:

- when the two singleton and geminate groups are merged the following values of r_s are obtained: V1d vs C1d r_s = - 0.471; V1d vs C2d r_s = - 0.474

- when the two utterance groups are kept separate, then the following r_s values are obtained: V1d vs C1d: r_s = - 0.271 for singletons and r_s not significative for geminates; V1d vs C2d: r_s = - 0.432 for singletons and r_s = - 0.321 for geminates.

Therefore, the correlation between V1d and C1d can be attributed to the presence of gemination. The same can be said for the correlation between V1d and C2d, although not with the same strength.

On the basis of the results on time-domain quantities, it was decided to use the parameters V1d, C1d, C2d, Cd, Cd/Utd, Cd/V1d, C1d/V1d and C2d/V1d to perform utterance classification based on time parameters using the Maximum Likelihood Criterion (MLC) (Dillon W.R. and Goldstein M., 1984). Results are reported in Table IV. The error percentages are: on Cd/Utd: 19.9%, on Cd/V1d: 16.7%, on C1d/V1d: 13.9%, on C2d/V1d: 30.3%, on V1d: 23.4%, on C1d: 16.9%, on C2d: 36.3%, on Cd: 17.6%. The best result is obtained using C1d/V1d but, if we consider specific consonants, the result is even better (for example, with C1d/V1d there are no mistakes for [ʤ] and with Cd there is only one mistake (percentage error 4.63%) for [ʧ]).

	Cd/Utd			Cd/V1d			C1d/V1d			C2d/V1d		
	EPP	Errors	Err. %	EPP	Errors	Err. %	EPP	Errors	Err. %	EPP	Errors	Err. %
Overall	0.43	86	19.91	1.6	72	16.67	0.76	60	13.89	0.78	131	30.32
Male	0.42	40	18.52	1.68	30	13.89	0.69	21	9.72	0.78	66	30.56
Female	0.43	45	20.83	1.55	40	18.52	0.85	36	16.67	0.78	65	30.09
[ʧ]	0.44	15	13.89	1.6	9	8.33	0.65	9	8.33	0.87	20	18.52
[ʤ]	0.37	6	5.56	1.26	3	2.78	0.91	0	0.00	0.36	17	15.74
[ts]	0.49	31	28.70	1.9	26	24.07	0.77	29	26.85	1.41	28	25.93
[dz]	0.42	15	13.89	1.45	14	12.96	0.76	13	12.04	0.75	18	16.67
[a]	0.43	23	15.97	1.29	19	13.19	0.74	15	10.42	0.78	38	26.39
[i]	0.48	32	22.22	1.75	25	17.36	0.92	22	15.28	1.13	41	28.47
[u]	0.44	27	18.75	1.57	21	14.58	0.68	18	12.50	0.87	41	28.47
	V1d			C1d			C2d			Cd		
	EPP	Errors	Err. %	EPP	Errors	Err. %	EPP	Errors	Err. %	EPP	Errors	Err. %
Overall	136	101	23.38	102	73	16.90	129	157	36.34	208	76	17.59
Male	135	47	21.76	93	26	12.04	146	82	37.96	215	36	16.67
Female	135	54	25.00	112	41	18.98	129	67	31.02	208	37	17.13
[ʧ]	137	20	18.52	95	10	9.26	139	30	27.78	224	5	4.63
[ʤ]	136	15	13.89	125	5	4.63	67	34	31.48	166	8	7.41

[ts]	124	34	31.48	91	27	25.00	157	28	25.93	244	19	17.59
[dz]	153	28	25.93	100	17	15.74	108	28	25.93	201	11	10.19
[a]	145	32	22.22	104	17	11.81	107	53	36.81	198	23	15.97
[i]	122	33	22.92	102	24	16.67	139	48	33.33	235	28	19.44
[u]	135	27	18.75	94	24	16.67	129	54	37.50	207	24	16.67

TABLE IV Results of the one-dimensional Maximum Likelihood Criterion used for classifying singleton vs. geminate utterances.

2.2 Results in the frequency domain

Table V reports the average frequency parameter values and their standard deviations obtained by working out the average of all utterances. All single measured values can be found in (Faluschi, 2000). Statistical analysis, using a multivariate ANOVA, indicates that there are not statistically meaningful differences between measurements, except for slight changes in the fundamental frequency (F0) when it is measured in very specific frames. F0 is about 14 Hz and 12 Hz higher in the geminate form (+9% and +8%) in the V1 offset, and V1 to C transition frames, respectively. No significant variations were observed in F1, F2 and F3. Formant amplitudes A1, A2 and A3 are significantly higher in their geminate form (1-3 dB) in V1 center, V1 offset, V1 to C transition and V2 onset frames. Performing a Maximum Likelihood Criterion based on the statistically significative frequency parameters, we are led to about 64% of correct classifications, an average which is quite low.

	F0	A0	F1	A1	F2	A2	F3	A3
V1 CENTER								
Singleton	160	13	550	34	1581	32	2895	26
(StD)	44	6	310	9	777	8	466	11
Geminate	169	13	552	36	1588	34	2870	26
(StD)	43	7	303	8	769	8	548	11
V1 OFFSET								
Singleton	151	13	487	29	1684	27	2877	23
(StD)	44	5	242	8	679	9	469	9
Geminate	165	14	506	32	1694	30	2792	25
(StD)	43	7	256	7	683	8	656	10
V1 TO C TRANSITION								
Singleton	148	12	448	26	1716	22	2864	21
(StD)	43	5	201	8	638	8	469	7
Geminate	160	13	466	27	1728	25	2778	22
(StD)	44	5	220	7	651	8	648	9

	F0	A0	F0	A0	F0	A0	F0	A0
C1 ONSET / C1 CENTER / C2 CENTER / C2 OFFSET								
Singleton	139	11	126	9	112	9	120	10
(StD)	36	5	32	5	23	4	27	5
Geminate	70	5	62	4	40	3	48	4
(StD)	78	7	69	6	58	6	64	7

	F0	A0	F1	A1	F2	A2	F3	A3
V2 ONSET								
Singleton	134	10	413	28	1724	25	2758	22

	F0	A0	F1	A1	F2	A2	F3	A3
(StD)	*32*	*5*	*167*	*6*	*508*	*7*	*409*	*7*
Geminate	138	11	415	29	1743	27	2689	23
(StD)	*33*	*5*	*162*	*7*	*488*	*7*	*614*	*8*
V2 CENTER								
	F0	**A0**	**F1**	**A1**	**F2**	**A2**	**F3**	**A3**
Singleton	130	10	476	30	1617	27	2798	21
(StD)	*31*	*5*	*258*	*6*	*653*	*7*	*421*	*8*
Geminate	133	10	480	31	1626	28	2724	21
(StD)	*32*	*5*	*250*	*7*	*644*	*7*	*627*	*9*

TABLE V Frequency-based parameters. Average values and Standard Deviation with respect to all repetitions, speakers, vowels and consonants. Frequencies are in Hz, amplitudes in dB.

2.3 Results in the energy domain

Table VI reports the data on the energy-based parameters. An ANOVA mulivariate test was performed in order to determine if statistically significative differences between averages exist. Bold characters indicate significantly different values. The general tendency is to pronounce geminate words with more emphasis than singleton ones. This is confirmed also by previous analysis where there is a higher amplitude of formants in geminate utterances. Performing a Maximum Likelihood Criterion based on the energy parameters, leads to a poor average 61.7% of correct classification.

	EtotV1	PmV1	EtotC1	PmC1	EtotC2	PmC2	EtotC	PmC
Singleton	93.7	**62.1**	75.9	**47.1**	**78.4**	**49.1**	**81.4**	59.1
(StD)	*6.0*	*5.5*	*5.9*	*6.0*	*4.8*	*5.3*	*4.1*	*4.5*
Geminate	94.1	**63.7**	76.9	**45.8**	**80.9** 5.1	**50.5**	**83.4**	59.4
(StD)	*5.7*	*5.2*	*6.0*	*5.9*		*5.5*	*4.5*	*4.7*
	EiV1cent	**EiV1-C1**	**EiC1cent**	**EiC1-C2**	**EiC2cent**	**EiC2offs**		
Singleton	**86.8**	**79.4**	**67.3**	68.6	**73.1**	**72.5**		
(StD)	*5.9*	*5.1*	*9.0*	*6.6*	*6.2*	*5.1*		
Geminate	**88.8**	**80.9**	**63.2**	68.2	**74.3**	**74.5**		
(StD)	*5.3*	*4.7*	*9.5*	*6.2*	*6.6*	*5.5*		

TABLE VI Energy-based parameters. Average values and Standard Deviation with respect to all repetitions, speakers, vowels and consonants. All values are in dB. Bold characters indicate significantly different values.

SECTION 3

3 Discussion and conclusion

Affricate consonants were analyzed in the present study. These consonants have peculiar and very distinctive characteristics that made necessary a split of the consonant itself into two parts: the first is named C1 and indicates the stop phase, while the second, named C2, indicates the fricative phase. This distinction had not been necessary in previous studies about other Italian consonants.

Overall, results of the present study on affricates confirm the observations of the above studies on the other classes of Italian consonants, i.e. there is a significant increase of consonant duration in geminates, and a corresponding reduction of the pre-consonant vowel length. These results are confirmed by previous studies concerning the perception of stops (Esposito and Di Benedetto, 1999) and nasals (Mattei and Di Benedetto, 2000).

In the time domain, there is evidence for a strong connection between gemination and phoneme durations. In particular, as it was for previous gemination studies on Italian, it was evident that when moving from singleton to geminate utterances, an increase of consonant duration (both in the C1 stop phase and C2 fricative phase) and a decrease of pre-consonant vowel duration were observed. There are no significant variations in second vowel (V2) duration. Even the total duration of the utterance (Utd) moves up in the geminate form, although not in such a significant way as in V1, C1 and C2. This observation leads to suppose that a compensation effect (even if incomplete) takes place between V1 and C durations. This compensation effect is the element that tends to make the phonetic rhythm almost unchanged.

In the frequency domain, F0 increases passing from singleton to geminate only in two frames, by an average quantity of about 13 Hz that is perceptively relevant. The amplitudes of the three measured formants , A1, A2 and A3 in each analysis frame, except for V2 center, were also measured. These three amplitudes are higher of about 2 dB average in geminate utterances compared to singleton ones.

As far as the energy and power of pronunciations are concerned, we see the tendency to give more emphasis to geminates compared to singletons.

For affricates, it was found that the average difference between singles and geminates in terms of V1d is 38 ms ($\approx$-25% for geminates), in C1d is 51 ms ($\approx$+62% for geminates), in C2d is 27 ms ($\approx$+28% for geminates) while considering the whole consonant Cd=C1d+C2d the difference is 78ms ($\approx$+44% for geminates). For fricatives (Giovanardi M., 1998) the average V1d difference was 49 ms ($\approx$-28% for geminates) and in Cd 98 ms ($\approx$+73% for geminates). As regards stops (Esposito A., Di Benedetto M.G., 1999), V1d difference was 43 ms ($\approx$-26% for geminates) and 92 ms for the stop closure duration (+101% for geminates). Finally, for nasals (Mattei M., Di Benedetto M.G., 1999) the average V1d difference was 59 ms ($\approx$-32% for geminates) and in Cd 121 ms ($\approx$+134% for geminates).

The straightforward comparisons that can made with previous studies are between the C1 occlusive phase of the affricate consonants and the stop consonants (Esposito A., Di Benedetto M.G., 1999) and also between the C2 fricative phase of the affricate consonants and the fricatives (Giovanardi M., 1998).

There is evidence to assert that differences between singleton and geminate durations are less sharp and clear than in other Italian consonants.

The singleton/geminate classification based on the Maximum Likelihood Criterion using the Cd and Cd/V1d parameters leads to the following results, belonging to previous studies:

1) 0.47% of errors in nasals for both parameters

2) 12 % of errors in fricatives for both parameters

3) 4% and 8% of errors in stops for Cd and Cd/V1d, respectively

In Table IV, it is possible to see that error percentages for affricates are worse compared to those obtained for other classes of consonants.

Finally, an additional similarity between nasals, fricatives, and stops was found; the Spearman Rank correlation coefficient between Cd and V1d was for all these consonants equal to values from -0.71 to -0.78, while for affricates this value is lower.

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

The general tendency of shortening the pre-consonant vowel and of lengthening the consonant in geminate utterance, observed on stops, fricatives and nasals in previous studies, is confirmed for affricates.

In previous studies, significant variations in the energies of utterances have not been observed, except for the total energy of the consonant in nasals which was stronger in the geminates; this effect was explained by the longer duration of the consonant since consonant power was constant. For affricates, the tendency is to emphasize both energy and power of the geminate utterance (even if differences are limited to very few dB). This effect seems peculiar to the affricate class.

Finally, we wish to point out that there are studies in progress concerning synthetised utterances of affricate consonants obtained using the articulatory synthetiser HLsyn (HLsyn Reference Manual, 1997); the project will pilot an experiment of perceptive analysis in order to support the idea that, as there is evidence in this study, the duration of phonemes is the most important factor to distinguish a singleton utterance from a geminate one. In order to illustrate this theory we insert below two link-buttons: the first connects to the singleton synthetised utterance of the [tʃ] articulated with the [a] vowel, while the second connects to the corresponding geminate synthetised utterance.

[atʃa](#)

[attʃa](#)

References

Abramson A. S. (1999) "Fundamental frequency as a cue to word-initial consonant length: Pattani Malay" ICPhS99 San Francisco pp 591-594.

Argiolas F., Macrì F., Di Benedetto M.G. "Acoustic analysis of Italian [r] and [l]," Journal of the Acoustical Society of America 97, no. 5, pt.2, pp.3418, 1995.

Arvaniti A. (1999) "Effects of speaking rate on the timing of single and geminate sonorants" ICPhS99 San Francisco pp 599-602.

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

Canepari L. (1979) "Introduzione alla fonetica" Piccola Biblioteca Einaudi.

Cohn A.C., Ham W.H., Podesva R. J.(1999) "The phonetic realization of singleton-geminate contrasts in three languages of Indonesia" ICPhS99 San Francisco pp 587-590.

Esposito A., Di Benedetto M.G (1999), "Acoustic and Perceptual Study of Gemination in Italian Stops", Journal of the Acoustical Society of America, completare.

Faluschi S. (2000) "Analisi Acustica e Sintesi delle consonanti affricate singole e geminate in Italiano" Degree Thesis in Electrical Engineering, Rome University "La Sapienza".

Giovanardi M. (1998) "Analisi Acustica e Sintesi delle consonanti fricative singole e geminate in Italiano" Degree Thesis in Electrical Engineering, Rome University "La Sapienza".

Giovanardi M. and Di Benedetto M.G. (1998)."Acoustic analysis of singleton and geminate fricatives in Italian" European student journal of language and speech.

HLsyn Reference Manual (1997)

Local J.and Simpson A.P. (1999) "Phonetic implementation of geminates in Malayalam nouns" ICPhS99 San Francisco pp 592-595.

Louali N.and Maddieson I. (1999) "Phonological contrast and phonetic realization: the case of Berber stops" ICPhS99 San Francisco pp 603-606.

Mattei M. and Di Benedetto M.G. (2000) "Acoustic analysis of singleton and geminate nasals in Italian" European student journal of language and speech.

Mattei M. (1999) "Analisi Acustica delle consonanti nasali singole e geminate in Italiano" Degree Thesis in Electrical Engineering, Rome University "La Sapienza".

Muljacic Z. (1972). "Fonologia della Lingua Italiana", (Il Mulino, Bologna).

Rochet,L.B., and Rochet,A.P. (1995) "The perception of the single-geminate consonant contrast by native speakers of Italian and Anglophones" in proceedings of ICPhS95 pp. 616-619.

Rossetti R. (1993) "Caratteristiche acustiche del fenomeno di geminazione nelle consonanti occlusive Italiane: applicazione all'adattamento automatico di pronunce straniere" Degree Thesis in Electrical Engineering, Rome University "La Sapienza".

Rossetti R. (1994) "Gemination of Italian stops", Journal of the Acoustical Society of America, 95, 2pSP25, pp.2874.

Shrotriya N., Siva Sarma A.S., Verma R., Agrawal S.S. (1995). "Acoustic and perceptual characteristics of geminate Hindi stop consonants", in Proceedings of ICPhS9, pp.132-135.

Vecsys (1989). The Unice User Manual (Vecsys - Chemin du Chene rond - 91570 Bièvres, France).