Consonant gemination in Italian: the nasal and liquid case

Maria Gabriella Di Benedetto (*), Luca De Nardis (**)

(*) Radcliffe Institute for Advanced Study at Harvard University, Cambridge, USA, and Sapienza University of Rome, Rome, Italy

(**) DIET Department, Sapienza University of Rome Rome, Italy

{mariagabriella.dibenedetto, luca.denardis}@uniroma1.it

Abstract

All Italian consonants affected by gemination, that are affricates, fricatives, liquids, nasals and stops were analysed within a project named GEMMA that lasted over a span of about 25 years. Results of the analysis on stops, as published in a previous paper, had shown that the main acoustic cue to gemination in Italian is closure duration, while frequency and energy domain parameters were not significantly affected by gemination. This paper - the first of a set of two covering all remaining consonants - addresses nasals and liquids; its companion paper addresses affricates and fricatives. Results on nasals and liquids confirm the findings on stops, and reinforce the hypothesis that the primary acoustic cue to gemination in Italian is durational in nature, and corresponds to a lengthened consonant duration. Results also show an inverse correlation between consonant and pre-consonant vowel durations. This correlation is, however, also present when considering singleton vs. geminate word sets separately, indicating a sort of duration compensation between these phonemes to eventually preserve rhythmical structures; this inverse correlation is reinforced when considering singleton and geminate sets combined. Classification tests of single vs. geminate consonants show that, for both nasals and liquids, best classification scores are obtained using consonant duration, that is a durational parameter. Although slightly less performing, the ratio between consonant and pre-consonant vowel durations is also a potential good candidate for automatic classification of nasals and liquids geminate vs singleton consonants in Italian.

1. 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 gemination is contrastive and therefore several words change their meaning as a function of the presence or absence of gemination of one consonant in the word. Gemination in Italian appears most often in disyllabic words forming minimal pairs, where lexical stress is placed on the first syllable of the word. Words belonging to minimal pairs are orthographically distinguished by a double grapheme of the geminate consonant (for example: papa (pope) vs. pappa (baby food), or casa (house) vs. cassa (box)). Native Italian speakers have a natural attitude in producing disyllabic words of minimal pairs identified by the presence or absence of consonant gemination. In Italian, moreover, gemination can be also observed across neighbour words of a same sentence, giving rise to a phenomenon that is peculiar of the Italian language, called “raddoppiamento sintattico.”

The identification of acoustic correlates of gemination in Italian, and the verification of their perceptual relevance, is a longstanding research challenge. A pioneering study addressed Italian stops (Rossetti, 1993; Rossetti, 1994; Esposito and Di Benedetto, 1999), based on the analysis of speech materials consisting in VCV vs. VCCV words. Results showed that consonant closure duration and pre-consonant vowel duration were affected by gemination. In particular, when gemination was present, the pre-consonant vowel was shortened, while consonant closure duration increased, leading to a suggestion that speakers may tend to preserve the rhythmic structure of the word. Similar observations were also reported by (Rochet and Rochet, 1995) and (Pickett et al., 1999), where the latter also observed some kind of constancy in the phenomenon across speaking rates. Evidence for relational acoustic relevance was also found in Japanese (Hirata and Whiton, 2005). Gemination was investigated at large in several other languages beyond Italian; evidence for consonant duration as the main acoustic cue to gemination was also found in stops and fricatives in Lebanese (Al-Tamimi and Khattab, 2011; Khattab and Al-Tamimi, 2014; Al-Tamimi and Khattab, 2015), in Hindi (Shrotriya et al., 1995),
in Cypriot Greek (Arvaniti, 1999; Arvaniti and Tserdanelis, 2000; Tserdanelis and Arvaniti, 2000), in Persian stops (Hansen, 2004), in three languages of Indonesia (Cohn et al., 1999), in Swedish and Iraqi Arabic (Hassan, 2003), in Japanese stops (Idemaru and Guion, 2008) and in Berber (Louali and Maddieson, 1999; Ridouane, 2007), although in Berber geminate stops lack their singleton counterpart. Other secondary acoustic cues such as shortening of the pre-consonant vowel in the presence of gemination were found in Italian (Esposito and Di Benedetto, 1999) and in Berber (Ridouane, 2007), but not in Persian (Hansen, 2004) nor in Arabic (Hassan, 2003).

Acoustic cues related to the frequency domain - rather than durational patterns - were found in an Austronesian language, Pattani Malay (Abramson, 1998). The study of Pattani Malay focused on the analysis of fundamental frequency (F0) variations with gemination of word-initial consonants. Findings were that F0 varied with gemination, although not for all consonantal classes. In particular, F0 in nasal consonants was not affected by gemination, while the opposite was true for stops, as also confirmed in a perceptual experiment (Abramson, 1999).

A Dravidian language, Malayalam (Local and Simpson, 1999), stands somewhat apart from others, for spectral and temporal properties seem to be equally relevant in characterizing gemination.

The speech group at Sapienza University of Rome, Italy, has been active in tackling the problem of finding acoustic cues to gemination in the Italian language during the last three decades; the Gemination project GEMMA (Di Benedetto, 2000; GEMMA, 2019) started at Sapienza in 1992, with the ambition of analyzing gemination for all Italian consonants occurring in both singleton and geminate forms. The analyzed consonants were stops, liquids, fricatives, nasals, and affricates. The first extensive publication output of the GEMMA project addressed gemination in stop consonants (Esposito and Di Benedetto, 1999).

Beyond stops, all published materials appeared either in abstracts or in currently out-of-print journals; results for liquids were presented in a meeting of the Acoustical Society of America (Argiolas et al., 1995), while for all other consonants in the former and no longer available (since mid-2011) copyright-free web journal named European Student Journal of Language and Speech “WEB-SLS” (fricatives: Giovanardi and Di Benedetto, 1998; nasals: Mattei and Di Benedetto, 2000; affricates: Faluschi and Di Benedetto 2000).

The purpose of this manuscript, and of its companion paper (Di Benedetto and De Nardis, 2019), is to extend, revisit and contextualize the work originally presented in the above “beyond stops” papers. Novel contributions of the present submissions include exhaustive statistical analyses on time, frequency and energy domain parameters, and the analysis of time domain parameters as potential test variables for the classification of singleton vs. geminate words. As a result, this paper and the companion paper provide a comprehensive assessment of gemination in Italian, and offer to the speech research community the benefit of accessing to results and data that do not appear anymore in a published form.

The reference value of the material is reinforced by providing, as integral part of this revisit, full access to the entire database on which the study has been founded. This database is a unique case of Italian consonants in VCV vs. VCCV words. A detailed description of the database is provided in Section 2, along with details on speech material for nasals and liquids, analyzed in the present paper. Acoustic analyses and statistical tests are presented in Section 3. Results of acoustic analysis are reported in Section 4. Section 5 provides a discussion and comparison of results for nasals vs. liquids, as well as the results of classification tests for singleton vs. geminate words. Section 6 draws conclusions and highlights future avenues of research.

2. Speech materials
2.1 The GEMMA database

The speech materials analyzed in the present work, and in its companion paper (Di Benedetto and De Nardis, 2019), are part of the GEMMA project database (GEMMA, 2019). This database is composed of disyllabic words, i.e. vowel–consonant–vowel (VCV) in the nongeminate case and vowel–consonant–consonant–vowel (VCCV) in the geminate case. The consonants in the words are stops (/b/, /d/, /g/, /p/, /t/, /k/), affricates (/tʃ/, /dʒ/, /ts/, /dz/), fricatives (/ʃ/, /s/, /z/), nasals (/m/, /n/) and liquids (/l/, /r/), that is all consonants that in the Italian language are generally accepted as appearing in either single or geminate forms in intervocalic position. The case of affricates is however a tricky and debatable one, as will be further discussed in the companion paper. The vowels in the words are /a, i, u/, that is a subset of Italian vowels /a, e, ɛ, i, o, ɔ, u/. Words are symmetric with respect to vowel.

Six adult Italian native speakers, three men and three women aged from twenty-four to fifty, participated in the recordings. The speakers were pronunciation defectless and free of dialectal inflexions.

The words in the GEMMA database were pronounced in isolation and not in carrier sentences, in order to limit the effect of factors such as intonation (Rossetti, 1993, 1994).
Words were written on cards that were presented to the speaker by the operator. Cards were shuffled after each recording session. The speech materials of the GEMMA database were recorded in the Speech Laboratory of the INFOCOM Department (now DIET Department) at the University of Rome 'La Sapienza' (Italy) using professional equipment, in a sound-treated room under the supervision of an acoustically trained subject. The entire set of words was recorded three times in three different recording sessions, leading to three repetitions for each word and for each speaker. In case of evident mispronunciations, the speaker was compelled to repeat the word.

The words were then digitized using the UNICE software produced by Vecsys (Vecsys, 2019). Speech signals were filtered at 5 kHz, sampled at 10 kHz, and each sample was quantized with 16 bits. Each signal was then stored by UNICE as a .sig file containing the samples, and a companion .key file containing information on sampling rate and quantization.

The GEMMA database is now available under an open source Creative Commons license; the original UNICE file doublets describing each speech signal were converted into .wav files using the sox open source utility, in order to offer a wide access to the material (GEMMA, 2019). The top folder of the database contains a README file providing a detailed description of its organization, briefly summarized as follows. The database is organized in five folders, one for each family of consonants: folder “Affricates” for affricates, folder “Fricatives” for fricatives, folder “Liquids” for liquids, folder “Nasals” for nasals and folder “Stops” for stops. Each of the above folders is further organized into six folders, one for each speaker, named “FS1”, “FS2”, “FS3”, for the three female speakers, and “MS1”, “MS2”, “MS3” for the three male speakers. The initials of the six speakers are stored in the README file. Each speaker folder contains the files for the three repetitions for that specific consonant set; the generic file name is in the form “<Word><Repetition><Speaker>.wav,” e.g., the first repetition for the word “iffl” for the first female speaker is named “IFF1FS1.wav”.

2.1. Nasals and liquids speech materials
In the Italian language, the set of nasal consonants that is generally accepted as appearing in both singleton and geminate forms is /m, n/, since /n/ appears only in the geminate form (Muljacic, 1972). Table I shows the set of words in the database containing nasal consonants; the consonants in the geminated form are transcribed as /mn/ and /nn/. Given the number of speakers (6 speakers), the number of repetitions (3 repetitions), the number of symmetrical vowel contexts (3 vowel contexts), the number of consonants (2 consonants) and the forms (singleton vs. geminate), a total of 6x3x3x2x2=216 words were recorded.

| a | ama | amma | an | ann |
|---|-----|------|---|-----|
| i | ini | immi | ini | inni |
| u | umu | ummu | unu | unnu |

Table I - Set of words of the GEMMA database that contain nasal consonants. Singleton consonants are indicated by /m, n/, while geminate consonants are indicated by /mn/ and /nn/.

Regarding liquids, the set of liquids that appear in the Italian language in both singleton and geminate forms is /l, r/ (Muljacic, 1972). Table II shows the set of words in the GEMMA database containing liquid consonants, where consonants in the geminated form are again represented by a double grapheme of the consonant. Given the number of speakers (6 speakers), the number of repetitions (3 repetitions), the number of symmetrical vowel contexts (3 vowel contexts), the number of consonants (2 consonants) and the forms (singleton vs. geminate), a total of 6x3x3x2x2=216 words were recorded.

| a | ala | alla | ara | arra |
|---|-----|------|-----|------|
| i | ili | illi | iri | irri |
| u | ulu | ullu | uru | urru |

Table II - Set of words in the GEMMA database containing liquid consonants. Singleton consonants are indicated by /l, r/, while geminate consonants are indicated by /ll, rr/.
3. Measurements and statistical tests

The analysed parameters refer to time, frequency, and energy domains. Measurements of the parameters were taken at specific times and specific frames that are defined in the Section 3.1.1. Time domain parameters are described in Section 3.1.2. Frequency domain and energy domain parameters are described in Sections 3.1.3 and 3.1.4, respectively. Finally, Section 3.1.5 describes the statistical tests that have been adopted to analyse the statistical significance of the variations of the parameters.

3.1.1. Reference times and reference frames

The analysed parameters were measured at specific instants in time, called reference times, that correspond to relevant acoustic events within the word. The identification of reference times was made based on the specific characteristics of each consonant. Reference times can be listed as follows (see Fig. 1):

- **Vowel 1 onset time (V1 onset)** – The pre-consonant vowel onset time, V1 onset, was identified by the appearance of a glottal pulse followed by other regular glottal pulses.
- **Vowel 1 offset time (V1 offset)** – The pre-consonant vowel offset time, V1 offset, was identified as the time at which glottal pulses disappear.
- **Vowel 2 onset time (V2 onset)** – The post-consonant vowel onset time, V2 onset, was identified by visual inspection of waveforms and spectrograms as the time instant at which a glottal pulse appeared, and an abrupt shift in formants was detected. The decision was also supported in specific cases by a short-term energy analysis and in some few cases by a perception test.
- **Vowel 2 offset time (V2 offset)** – The post-consonant vowel offset time, V2 offset, was typically matched with the disappearance of the second and higher formants. In specific cases, mostly in words including the [u] vowel, this reference time was set as the time at which the amplitude of the signal decreased below 90% of its peak value.
- **Consonant onset time (C onset)** – The consonant onset time, C onset, coincides with V1 offset.

![Figure 1 - Reference times selected for the computation of the acoustic parameters: V1 onset: reference time corresponding to onset of pre-consonant vowel; V1 offset: offset of pre-consonant vowel, corresponding to onset of the consonant \( C_\text{onset} \); V2 onset: onset of post-consonant vowel, corresponding to the offset of the consonant, referred to as \( C_\text{offset} \); V2 offset: offset of post-consonant vowel.](image-url)
• Consonant offset (C_{offset}) – The consonant offset time C_{offset} coincides with V_{2 onset}.

A set of reference frames, each consisting of 256 samples, was also defined, with respect to reference times. Figure 2 shows the reference frames, that are defined as follows:

- **V1 CENTRE** – frame located at V1 center, i.e. centered on \( \frac{V_{1\text{onset}} + V_{1\text{offset}}}{2} \);
- **V1 OFFSET** – frame located at the offset of V1, right before V_{1 offset};
- **V1-TO-C TRANSITION** – frame located at the transition between V1 and C, centered on V_{1 offset};
- **C ONSET** – frame located at the onset of the consonant, i.e. starting at V_{1 offset};
- **C CENTRE** – frame located at C center, i.e. centered on \( \frac{V_{1\text{onset}} + C_{offset}}{2} \);
- **C OFFSET** – frame located at the offset of the consonant, i.e. ending at C_{offset};
- **V2 ONSET** – frame located at the onset of V1, i.e. starting at V_{2 onset};
- **V2 CENTRE** – frame located at the center of V2, i.e. centered on \( \frac{V_{2\text{onset}} + V_{2\text{offset}}}{2} \).

![Reference frames definition](image)

**Figure 2** – Reference frames defined with respect to the reference times introduced in Figure 1. Each reference frame contains 256 samples.

### 3.1.2. Time domain parameters

Figure 3 shows the time domain parameters, defined as follows:

- duration of the pre-consonant vowel V1d, defined as V1d=V_{1 offset}-V_{1 onset};
- duration of the consonant Cd, defined as Cd=C_{offset}-C_{onset};
- duration of the post-consonant vowel V2d, defined as V2d=V_{2 offset}-V_{2 onset};
- duration of the entire word Utd, defined as Utd=V_{2 offset}-V_{1 onset}.

![Time domain parameters](image)

**Figure 3** – Time domain parameters defined with respect to reference times, as introduced in Figure 1: V1d, duration of first vowel; Cd, duration of consonant; V2d, duration of second vowel; Utd, duration of the entire word.

### 3.1.3. Frequency domain parameters

In order to carry out the analysis in the frequency domain speech signals were pre-emphasized with a pre-emphasizing filter with \( \alpha=0.95 \) and windowed using a Hamming window of 256 samples. Spectrograms, DFT (Discrete Fourier Transform) and LPC (Linear Predictive Coding) spectra were examined and compared to extract the following parameters:

- Fundamental frequency F0;
- First three formant frequencies F1, F2 and F3.
The above parameters were evaluated with respect to the reference frames as follows (see Figure 2 for reference):

- V1 CENTRE: F0, F1, F2 and F3;
- V1 OFFSET: F0, F1, F2 and F3;
- V1-TO-C TRANSITION: F0, F1, F2 and F3;
- C ONSET: F0 (both nasals and liquids), F1, F2 and F3 (liquids only);
- C CENTRE: F0 (both nasals and liquids), F1, F2 and F3 (liquids only);
- C OFFSET: F0 (both nasals and liquids), F1, F2 and F3 (liquids only);
- V2 ONSET: F0, F1, F2 and F3;
- V2 CENTRE: F0, F1, F2 and F3.

3.1.4. Energy domain parameters
The following energy domain parameters were defined:

- total energy of V1, indicated as \( E_{\text{totV1}} \), defined as follows:
  \[
  E_{\text{totV1}} = \sum |X_i|^2, 
  \]
  where \( X_i \) is i-th sample falling in the time interval \([V1_{\text{onset}}, V1_{\text{offset}}]\), corresponding to the duration of V1;
- average power of V1, defined as follows:
  \[
  P_{\text{V1}} = E_{\text{totV1}} / N_{\text{V1}}, 
  \]
  where \( N_{\text{V1}} \) is the number of samples within the interval \([V1_{\text{onset}}, V1_{\text{offset}}]\);
- total energy of C, indicated as \( E_{\text{totC}} \) and computed as for V1, but over the interval \([C_{\text{onset}}, C_{\text{offset}}]\), corresponding to the duration of C;
- average power of C, indicated as \( P_C \), and computed from \( E_{\text{totC}} \) as for \( P_{\text{V1}} \), but dividing by the number of samples within the interval \([C_{\text{onset}}, C_{\text{offset}}]\);
- instantaneous energy at V1 CENTRE, indicated as \( E_{\text{V1}} \), and defined as:
  \[
  E_{\text{V1}} = \sum |X_i|^2, 
  \]
  where \( X_i \) is i-th sample belonging to the V1 CENTRE reference frame;
- instantaneous energy at the transition V1-TO-C, indicated as \( E_{\text{V1-C}} \), and computed as for \( E_{\text{V1}} \) but in the V1-TO-C TRANSITION reference frame;
- instantaneous energy at C CENTRE, indicated as \( E_{\text{C}} \), and computed as for \( E_{\text{V1}} \);
- instantaneous energy at C OFFSET, indicated as \( E_{\text{Coff}} \) and computed as for \( E_{\text{V1}} \).

All energy domain parameters listed above were expressed in logarithmic form \((10 \log_{10}(x))\).

3.1.5. Statistical tests
The following statistical tests were performed on the parameters defined in the previous subsections (Dillon W.R. and Goldstein M., 1984):

- Repeated measurements ANOVA and multi-factor univariate ANOVA, used to determine whether average values of the parameters presented statistically significant differences between different groups of words;
- Spearman Rank Correlation Coefficient, used to detect correlations between the different parameters;
- Maximum Likelihood Classification (MLC) test, used to determine which parameters could be used for classification of singleton vs. geminate words.

4. Results
This section presents results for time, frequency and energy domain parameters for both nasals and liquids averaged over repetitions.

4.1. Results on nasals
4.1.1. Results in the time domain
Table III shows the average values over repetitions and speakers of V1d, Cd, V2d and Utd for nasal consonants \([\text{m, n}]\), and corresponding standard deviations. Results highlight a general tendency to shorten the pre-consonant vowel duration V1d and lengthen consonant duration Cd in geminate vs. singleton words, while the post-consonant vowel duration V2d does not appear to be affected by gemination in a systematic form. Geminate words were in average – over all words - about 14% longer than singletons.
A repeated measurements ANOVA test was performed on female and male speakers data separately, averaged over repetitions. Form (singleton vs. geminate) was used as a between-subjects factor, while Vowel ([a, i, u]) and Consonant ([m, n]) were considered as within-subjects factors. For each parameter, Table IV contains the test variable F and the corresponding p value for each factor and for the interaction between each within-subjects factor and the between-subjects factor; bold values indicate significant values, with threshold set as p=0.05.

Table IV – Results of the repeated measurements multivariate ANOVA test performed on time domain parameters for words containing nasals on female and male speakers separately, averaging data over repetitions; test variable F and corresponding probability p at which the null hypothesis can be rejected are presented for the between-subjects factor Form (singleton vs. geminate), for the within-subjects factors Vowel ([a, i, u]) and Consonant ([m, n]), and for their interactions; bold characters indicate significantly different values, with threshold set as p<0.05.

Table IV shows that gemination has a significant impact on the average value of Cd and V1d for both female and male speakers, and of Utd for female speakers. No significant variations were observed for V2d. Vowel has a significant impact on the Cd parameter for both female and male speakers; the same behavior can be observed for V2d and Utd. As for the Consonant factor, significant variations can be observed for V1d and V2d for both female and male speakers, and for Utd for female speakers.
In order to get further insight on the impact of gemination, additional univariate ANOVA tests were carried out separately for each vowel and consonant, considering Form as the only fixed factor. Male and female speakers were in this case combined, since Table IV highlighted no major differences for the two genders with respect to gemination. Results are presented in Table V, showing the test variable F and corresponding probability p of validity of the null hypothesis; values in bold indicate statistically significant variations between singleton vs. geminate groups, with threshold set as p<0.05.

Results of Table V confirm that Cd and V1d are both strongly impacted by gemination; variations of both parameters between singletons and geminates groups were in fact significant for all combinations of consonants and vowels. A weaker significance was observed for Ut, with significant variations in all cases but with markedly larger p values. Finally, the post-consonant vowel duration V2d did not vary significantly between singletons vs. geminates for any combination of vowels and consonants.

|        | a   | i   | u   |
|--------|-----|-----|-----|
| m      |     |     |     |
| F(1,34) | 84.98 | 324.43 | 59.67 |
| p      | 8.98E-11 | 5.8E-19 | 0.4987 |
| n      | 114.91 | 377.91 | 2.43 |
| p      | 1.91E-12 | 5.42E-20 | 0.1284 |

Table V - Test variable F and corresponding probability p at which the null hypothesis can be rejected in the univariate ANOVA test performed on time domain parameters for words containing nasals using the Form (singleton vs. geminate) as fixed factor, for each combination of consonants [m, n] and vowels [a, i, u]; bold characters indicate significantly different values, with threshold set as p<0.05.

Next, a Spearman Rank correlation coefficient test was carried out in order to verify whether any correlation between time domain parameters could be identified in relation to gemination; results are presented in Table VIa) for singleton and geminated words separately, and in Table VIb) for all combined words.

|        | a   | i   | u   |
|--------|-----|-----|-----|
| m      |     |     |     |
| V1d s. | 1.00 | -0.15 | 0.45 |
| Cd s.  | -0.15 | 1.00 | -0.09 |
| V2d s. | 0.45 | -0.09 | 1.00 |
| V1d g. | 1.00 | -0.28 | 0.39 |
| Cd g.  | -0.28 | 1.00 | -0.15 |
| V2d g. | 0.39 | -0.15 | 1.00 |

Table VI - Spearman Rank Correlation Coefficient r, of time domain parameters for words containing singleton and geminate nasals (Table VIa)), and for all words, singleton and geminate combined (Table VIb)). Bold characters indicate significant correlations, with threshold set at p<0.05.

Note that correlation coefficients close to 0 indicate negligible correlation between parameters, positive coefficients indicate direct correlation, and negative coefficients indicate inverse correlation. Table VI shows that a strong inverse correlation is present for V1d vs. Cd in the combined group, while a weaker one can be observed for the group of geminated words; no correlation was observed for V1d vs. Cd for singleton words. All groups are characterized by a significant positive correlation between V1d and V2d suggesting that a weak compensation for the lengthening of the consonant involves V2d as well. However, Cd vs. V2d negative correlation was weaker, and only significant in the combined group.

### 4.1.2. Results in the frequency domain

Tables VII and VIII show the mean and standard deviation of frequency domain parameters, for female vs. male speakers, singleton vs. geminate forms, and for each vowel, in reference frames: 1) V1 CENTER, 2) V1 OFFSET,
3) V1-TO-C TRANSITION (Table VII) and 4) C ONSET, 5) C CENTER, 6) C OFFSET, 7) V2 ONSET, 8) V2 CENTER (Table VIII). Values in both tables are averaged over consonants, speakers and repetitions. Results show an increased F0 average in geminate words for male speakers, in all frames, while no clear effect of gemination can be observed on pitch for female speakers, and on formants for neither group of speakers.

| V1 CENTER | Female (Hz) | Male (Hz) |
|-----------|-------------|-----------|
|            | F0  | F1  | F2  | F3  | F0  | F1  | F2  | F3  |
| Singleton | Mean | 189 | 1000 | 1488 | 3064 | 114 | 835 | 1329 | 2571 |
|           | Std  | 43  | 100  | 71   | 157  | 6   | 29  | 70   | 201  |
| Geminate  | Mean | 188 | 1028 | 1582 | 3020 | 119 | 831 | 1348 | 2769 |
|           | Std  | 41  | 92   | 109  | 187  | 8   | 25  | 53   | 238  |
| Singleton | Mean | 197 | 390  | 2786 | 3565 | 127 | 277 | 2295 | 3220 |
|           | Std  | 44  | 97   | 174  | 390  | 8   | 27  | 28   | 142  |
| Geminate  | Mean | 201 | 399  | 2776 | 3559 | 134 | 280 | 2306 | 3237 |
|           | Std  | 41  | 84   | 150  | 423  | 7   | 7   | 25   | 112  |

| V1 OFFSET | Female (Hz) | Male (Hz) |
|-----------|-------------|-----------|
|            | F0  | F1  | F2  | F3  | F0  | F1  | F2  | F3  |
| Singleton | Mean | 178 | 913  | 1499 | 3098 | 108 | 790 | 1332 | 2536 |
|           | Std  | 43  | 55   | 153  | 210  | 6   | 38  | 131  | 208  |
| Geminate  | Mean | 185 | 983  | 1549 | 3048 | 118 | 809 | 1312 | 2693 |
|           | Std  | 42  | 65   | 161  | 216  | 10  | 28  | 142  | 289  |
| Singleton | Mean | 184 | 377  | 2776 | 3528 | 119 | 297 | 2317 | 3251 |
|           | Std  | 44  | 87   | 162  | 406  | 8   | 19  | 47   | 137  |
| Geminate  | Mean | 195 | 390  | 2769 | 3499 | 133 | 280 | 2306 | 3144 |
|           | Std  | 44  | 86   | 178  | 483  | 6   | 16  | 15   | 201  |
| Singleton | Mean | 182 | 368  | 801  | 2973 | 122 | 316 | 679  | 2382 |
|           | Std  | 79  | 158  | 99   | 182  | 7   | 27  | 133  | 102  |
| Geminate  | Mean | 199 | 405  | 783  | 2971 | 138 | 288 | 714  | 2347 |
|           | Std  | 40  | 80   | 158  | 80   | 9   | 13  | 149  | 71   |

| V1-TO-C TRANSITION | Female (Hz) | Male (Hz) |
|--------------------|-------------|-----------|
|                    | F0  | F1  | F2  | F3  | F0  | F1  | F2  | F3  |
| Singleton          | Mean | 178 | 883  | 1514 | 3083 | 107 | 850 | 1283 | 2532 |
|                    | Std  | 42  | 34   | 248  | 202  | 6   | 26  | 176  | 193  |
| Geminate           | Mean | 184 | 957  | 1519 | 3057 | 116 | 833 | 1278 | 2630 |
|                    | Std  | 41  | 61   | 217  | 220  | 10  | 47  | 150  | 310  |
| Singleton          | Mean | 181 | 355  | 2756 | 3524 | 116 | 295 | 2332 | 3251 |
|                    | Std  | 43  | 88   | 170  | 416  | 9   | 20  | 47   | 176  |
| Geminate           | Mean | 193 | 381  | 2745 | 3565 | 132 | 282 | 2317 | 3146 |
|                    | Std  | 43  | 88   | 217  | 505  | 7   | 16  | 48   | 174  |
| Singleton          | Mean | 178 | 345  | 790  | 2986 | 120 | 310 | 712  | 2427 |
|                    | Std  | 77  | 142  | 133  | 149  | 8   | 24  | 184  | 144  |
| Geminate           | Mean | 195 | 390  | 781  | 3011 | 135 | 282 | 714  | 2330 |
|                    | Std  | 40  | 72   | 113  | 154  | 10  | 18  | 154  | 120  |

**Table VII** – Average and standard deviation of pitch F0 and formants F1, F2 and F3 in reference frames V1 CENTER, V1 OFFSET and V1-TO-C TRANSITION for words containing nasals, for female vs. male speakers, averaged over repetitions, speakers and consonants (frequencies are in Hz).
A multi-factor univariate ANOVA test was carried using Form, Vowel and Consonant as fixed factors on female vs. male speakers. Results are presented in Table IX, that shows a factor vs. parameter matrix: a checked cell at the intersection between a factor and a parameter indicates a significant difference in the average value of the parameter due to that factor. Results in Table IX indicate that Form does not cause significant differences for any of the frequency domain parameter for female speakers, while, for male speakers, F0 shows a significant difference in the three frames related to the first vowel as well as in the C ONSET frame. Vowel induced, as expected, significant differences in both F0 (intrinsic pitch) for V1-related frames, and in formants F1, F2 and F3, in frames related to both V1 and V2. Factor Consonant led to significant differences only in sporadic cases, in
particular in frames V1 OFFSET, V1-TO-C TRANSITION, V2 ONSET and V2 CENTER, and only for formant F2.
Overall, in nasals frequency domain parameters do not seem to provide much information about gemination across speakers of different genders.

|                  | Female |         |         |         | Male  |         |         |         |
|------------------|--------|---------|---------|---------|-------|---------|---------|---------|
|                  | F0     | F1      | F2      | F3      | F0    | F1      | F2      | F3      |
| V1 CENTER        | Form   |         |         |         | X     |         |         |         |
|                  | Vowel  | X       | X       | X       | X     | X       | X       | X       |
|                  | Consonant |   |         |         |       |         |         |         |
| V1 OFFSET        | Form   |         |         |         | X     |         |         |         |
|                  | Vowel  | X       | X       | X       | X     | X       | X       | X       |
|                  | Consonant |   |         |         |       |         |         |         |
| V1-TO-C TRANSITION | Form | X       | X       | X       | X     | X       | X       | X       |
|                  | Vowel  |         |         |         |       |         |         |         |
|                  | Consonant |   |         |         |       |         |         |         |
| C ONSET          | Form   |         |         |         | N/A   | X       | N/A     | N/A     |
|                  | Vowel  |         |         |         |       | X       | N/A     | N/A     |
|                  | Consonant |   |         |         |       |         |         |         |
| C CENTER         | Form   | N/A     |         |         |       | N/A     |         |         |
|                  | Vowel  |         |         |         |       |         |         |         |
|                  | Consonant |   |         |         |       |         |         |         |
| C OFFSET         | Form   | N/A     |         |         |       | N/A     |         |         |
|                  | Vowel  |         |         |         |       |         |         |         |
|                  | Consonant |   |         |         |       |         |         |         |
| V2 ONSET         | Form   |         |         |         |       |         |         |         |
|                  | Vowel  | X       | X       | X       | X     | X       | X       | X       |
|                  | Consonant |   |         |         |       |         |         |         |
| V2 CENTER        | Form   |         |         |         |       |         |         |         |
|                  | Vowel  | X       | X       | X       | X     | X       | X       | X       |
|                  | Consonant |   |         |         |       |         |         |         |

Table IX – Results of the multi-factor univariate ANOVA test performed on frequency domain parameters in vowel reference frames V1 CENTER, V1 OFFSET, V1-TO-C TRANSITION, C ONSET, C CENTER, C OFFSET, V2 ONSET and V2 CENTER for words containing nasals using Form, Vowel and Consonant as fixed factors; a checked cell at the intersection between a parameter and a factor indicates a significant difference between average values for the parameter with respect to the factor.

4.1.3. Results in the energy domain
Table X shows the average values of energy domain parameters (for a list of parameters refer to Section 3.1.4). Since in the case of energy domain parameters the impact of gender was not expected to be as strong as for frequency domain parameters, results are presented here averaged over all speakers and repetitions. No clear trend can be observed from the data presented in Table X.
**Table X** - Average and standard deviation of energy domain parameters for each combination of consonants [m, n], vowels [a, i, u] and singleton vs. geminate form, averaged over repetitions and speakers (values are in logarithmic form; for a list of parameters refer to Section 3.1.4).

A multi-factor univariate ANOVA test was thus performed in order to determine if statistically significative differences between averages exist. The test considered the fixed factors Form, Vowel, Consonant and Gender, and was applied to all words combined. Results of the ANOVA test are presented in Table XI as a matrix of factor vs. parameter in which a checked cell indicates a significant difference in the average value of the parameter due to factor. Table IX shows that $E_{\text{mc}}$ varies significantly with Form (gemination). As for the other factors, Consonant led to significant differences for all parameters while Vowel led to significant variations for all parameters related to V1 except for $E_{\text{IV1-C}}$ (energy of transition frame from vowel to consonant). Finally, the Gender factor led to no significant variations.

| Consonant | E$_{\text{IV1}}$ | P$_{\text{IV1}}$ | E$_{\text{IV1C}}$ | P$_{\text{IV1C}}$ | E$_{\text{IV1Cent}}$ | E$_{\text{IV1-C}}$ | E$_{\text{Cent}}$ | E$_{\text{Offset}}$ |
|-----------|-----------------|-----------------|-----------------|-----------------|------------------|-----------------|-----------------|-----------------|
| a         | Mean            |                 |                 |                 |                  |                 |                 |                 |
|           | 102.00          | 70.06           | 89.06           | 59.61           | 94.50            | 88.33           | 83.50           | 84.28           |
|           | Std             | 4.51            | 4.70            | 5.23            | 5.44             | 4.75            | 5.24            | 5.58            | 5.27            |
| amma      | Mean            | 98.39           | 67.83           | 91.06           | 58.00            | 92.61           | 86.39           | 81.94           | 82.56           |
|           | Std             | 3.30            | 2.90            | 6.19            | 6.41             | 3.14            | 4.30            | 6.82            | 7.02            |
| ana       | Mean            | 100.11          | 67.00           | 87.56           | 58.67            | 92.00           | 84.78           | 82.56           | 82.89           |
|           | Std             | 2.75            | 2.94            | 5.27            | 5.95             | 2.98            | 4.54            | 6.16            | 6.01            |
| ania      | Mean            | 98.17           | 66.72           | 92.00           | 59.11            | 92.06           | 85.11           | 82.83           | 82.28           |
|           | Std             | 3.13            | 2.94            | 5.47            | 5.54             | 3.23            | 4.61            | 6.09            | 5.69            |
| i         | Mean            | 92.72           | 60.44           | 89.11           | 59.28            | 84.72           | 84.44           | 82.89           | 82.78           |
|           | Std             | 5.32            | 5.30            | 7.12            | 7.07             | 5.00            | 7.29            | 6.54            | 6.70            |
| immi      | Mean            | 91.83           | 61.39           | 93.67           | 60.11            | 85.72           | 85.33           | 83.67           | 83.78           |
|           | Std             | 3.92            | 3.89            | 5.93            | 6.44             | 3.94            | 4.48            | 6.92            | 7.65            |
| ini       | Mean            | 91.89           | 59.22           | 89.11           | 59.72            | 83.06           | 84.44           | 83.44           | 83.61           |
|           | Std             | 4.42            | 4.58            | 6.17            | 6.74             | 4.67            | 6.30            | 6.67            | 6.86            |
| inni      | Mean            | 92.11           | 61.67           | 94.11           | 60.56            | 85.61           | 84.78           | 84.28           | 84.28           |
|           | Std             | 4.52            | 4.09            | 5.61            | 5.91             | 4.64            | 4.45            | 6.97            | 6.93            |
| u         | Mean            | 94.56           | 62.11           | 87.11           | 57.11            | 87.17           | 82.11           | 80.94           | 81.94           |
|           | Std             | 2.84            | 2.92            | 5.75            | 6.20             | 2.71            | 5.56            | 6.62            | 6.37            |
| umnu      | Mean            | 94.67           | 63.61           | 91.06           | 58.11            | 88.89           | 82.61           | 82.00           | 82.78           |
|           | Std             | 3.98            | 3.79            | 4.98            | 5.52             | 4.08            | 5.14            | 5.68            | 6.99            |
| unnu      | Mean            | 94.72           | 61.67           | 88.78           | 59.28            | 86.39           | 84.44           | 83.17           | 83.33           |
|           | Std             | 4.29            | 4.58            | 6.41            | 6.72             | 4.55            | 6.35            | 6.72            | 7.14            |

| Form      | E$_{\text{IV1}}$ | P$_{\text{IV1}}$ | E$_{\text{IV1C}}$ | E$_{\text{IV1Cent}}$ | E$_{\text{IV1-C}}$ | E$_{\text{Cent}}$ | E$_{\text{Offset}}$ |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Vowel     | X               | X               | X               | X               | X               | X               | X               |
| Consonant | X               | X               | X               | X               | X               | X               | X               |
| Gender    | X               | X               | X               | X               | X               | X               | X               |

**Table XI** - Results of the multi-factor univariate ANOVA test performed on energy domain parameters using Form, Vowel, Consonant and Gender for all words containing nasals; a checked cell indicates a significant difference between average values for the parameter with respect to the factor.

### 4.2. Results on liquids

#### 4.2.1. Results in the time domain

The time domain parameters listed in Section 3.1.2 were computed for each of the 108 singleton and 108 geminate liquid words.
Results are reported in Table XII, that contains the average values and standard deviations of V1d, Cd, V2d and Utd for all combinations of vowels [a, i, u] and consonants [r, l] in geminate vs. singleton forms, averaged over all repetitions and speakers.

Table XII shows that, as regards V1d and Cd, liquids behave somewhat like nasals; V1d tends to decrease with gemination, while the opposite is true for Cd. No clear trend can be observed for V2d and Utd.

|        | V1d (msecs) | Cd (msecs) | V2d (msecs) | Utd (msecs) |
|--------|-------------|------------|-------------|-------------|
|        | Mean        | StD        | Mean        | StD         | Mean        | StD         |
| a      | 180.77      | 21.54      | 73.93       | 13.78       | 113.98      | 20.82       |
|        | 127.25      | 24.73      | 201.69      | 28.64       | 99.59       | 16.94       |
|        | 200.54      | 29.19      | 89.99       | 19.68       | 107.99      | 28.40       |
|        | 145.28      | 31.94      | 202.64      | 23.82       | 96.98       | 17.52       |
| i      | 169.74      | 22.87      | 81.69       | 13.81       | 112.74      | 21.38       |
|        | 127.61      | 17.51      | 204.96      | 27.17       | 112.56      | 23.97       |
|        | 191.79      | 151.89     | 95.32       | 18.22       | 115.11      | 17.03       |
|        | 25.28       | 28.17      | 203.36      | 31.43       | 110.84      | 17.77       |
| u      | 176.79      | 26.04      | 90.87       | 12.62       | 120.93      | 17.85       |
|        | 123.52      | 25.83      | 207.20      | 30.54       | 103.49      | 19.68       |
|        | 188.13      | 22.26      | 85.78       | 12.89       | 108.48      | 24.91       |
|        | 138.40      | 28.85      | 205.65      | 28.42       | 100.23      | 23.62       |

Table XII - Average values and standard deviations (in milliseconds) of V1d, Cd, V2d and Utd for words containing liquids, averaged over all repetitions and speakers.

Along the same approach adopted in Section 4.1.1 for nasals, a repeated measurements ANOVA test was performed on female and male speakers data separately, after averaging over repetitions, using Form (singleton vs. geminate) as a between-subjects factor, and Vowel [a, i, u] and Consonant [l, r] as within-subjects factors. Results of the test are presented in Table XIII. For each parameter, Table XIII shows the test variable F and the corresponding p value for each factor and for the interaction between each within-subjects factor and the between-subjects factor. Bold values indicate significant variations, with threshold set as p<0.05.

|        | Female | Male |
|--------|--------|------|
|        | F      | p    | F    | p    |
| V1d    | Form   | 3.483 | 0.135 | 154.8 | <0.001 |
|        | Vowel*Form | 2.393 | 0.153 | 0.252 | 0.783 |
|        | Consonant*Form | 0.348 | 0.587 | 0.213 | 0.669 |
|        | Vowel | 0.820 | 0.474 | 1.292 | 0.326 |
|        | Consonant | 5.200 | 0.085 | 42.250 | 0.003 |
| Cd     | Form   | 71.124 | 0.001 | 500.170 | <0.001 |
|        | Vowel*Form | 0.201 | 0.822 | 0.179 | 0.839 |
|        | Consonant*Form | 0.022 | 0.890 | 8.287 | 0.045 |
|        | Vowel | 0.211 | 0.814 | 1.155 | 0.353 |
|        | Consonant | 0.300 | 0.613 | 4.409 | 0.104 |
| V2d    | Form   | 1.498 | 0.288 | 0.175 | 0.698 |
|        | Vowel*Form | 0.977 | 0.417 | 2.256 | 0.167 |
|        | Consonant*Form | 1.046 | 0.364 | 0.098 | 0.770 |
|        | Vowel | 0.469 | 0.642 | 8.569 | 0.01 |
|        | Consonant | 0.049 | 0.836 | 6.005 | 0.07 |
| Utd    | Form   | 4.711 | 0.096 | 13.417 | 0.022 |
|        | Vowel*Form | 1.621 | 0.256 | 1.022 | 0.402 |
|        | Consonant*Form | 0.937 | 0.388 | 3.532 | 0.133 |
|        | Vowel | 0.001 | 0.999 | 5.079 | 0.038 |
|        | Consonant | 3.373 | 0.140 | 15.997 | 0.016 |

Table XIII – Results of the repeated measurements ANOVA test for liquids performed on time domain parameters, for female vs. male speakers separately. Values are averaged over all repetitions. Test variable F and corresponding probability p at which the null hypothesis can be rejected are presented for the between-subjects factor Form (singleton vs. geminate), for the within-subjects factors Vowel [a, i, u] and Consonant [l, r], and for their interactions. Bold characters indicate significant variations, with threshold set as p<0.05.
In terms of gemination, results in Table XIII highlight a significant variation of Cd for both female and male speakers, while only male speakers show a significant variation of both V1d and Utd. No significant variations were observed for V2d.

As for other factors, Consonant has a significant impact on V1d and Utd for male speakers. Finally, Vowel was significant only for V2d and male speakers.

As for nasals (see Section 4.1.1), additional univariate ANOVA tests for the Form factor (gemination) were carried out for each combination of vowel and consonant separately, on combined female and male speakers data. Results are presented in Table XIV, and confirm the combined results presented in Table XIII. Consonant duration Cd is strongly affected by gemination for all combinations of vowels and consonants. Gemination also has an impact on V1d in all cases, albeit with larger p values, and on Utd with an even weaker significance. As a side note, a significant variation for V2d was observed but only for [l] coarticulated with [a] and [u].

|       | a         | i         | u         |
|-------|-----------|-----------|-----------|
| F(1,34) | 47.95 | 5.18     | 38.53     | 43.97    | 37.98 | 223.05 | 7.75   | 11.08   |
| p     | 5.58E-08 | 0.0293 | 9.46E-07 | 0.98     | 1.32E-07| 5.3E-07| 0.0087 | 0.0021 |

Table XIV – Test variable F and corresponding probability p at which the null hypothesis can be rejected obtained in the univariate ANOVA test performed on time domain parameters for words containing liquids using Form (singleton vs. geminate) as fixed factor, for each combination of consonants [l, r] and vowels [a, i, u]. Bold characters indicate significantly different values, with threshold set as p<0.05.

Finally, the Spearman Rank Correlation Coefficient r was evaluated, for both singleton and geminate group, first separately and then combined. Results are presented in Table XVa) and Table XVb).

Table XV – Spearman Rank Correlation Coefficient r, of time domain parameters for singleton and geminate liquid words separately (Table XVa), and on all words combined (Table XVb). Bold characters indicate significant correlations, with threshold set at p<0.05.

V1d s.  V1d g.  Cd s.  Cd g.  V2d s.  V2d g.
Singleton  1.00  0.24  0.10  0.26  0.24   -0.26
Geminate  1.00  0.75  0.35  -0.28

V1d g.  Cd g.  V2d g.
Singleton  0.47  -0.28  1.00
Geminate  0.38  -0.32  1.00

4.2.2. Results in the frequency domain

The analysis in the frequency domain of liquids regarded frequency domain parameters F0, F1, F2 and F3 for both vowel related reference frames, and pitch F0 for consonant related frames, as defined in Section 3.1.1. The average value and standard deviations of F0, F1, F2 and F3 in V1 CENTER, V1 OFFSET and V1-TO-C TRANSITION reference frames are presented in Table XVI, while Table XVII presents average value and standard deviation of F0 in C ONSET, C CENTER and C OFFSET reference frames and of F0, F1, F2 and F3 in V2 ONSET and V2 CENTER reference frames. Data were obtained for female vs. male speakers separately and for
each combination of vowels [a, i, u] and forms (singleton vs. geminate), averaged over all speakers, consonants and repetitions.

|       | V1 CENTER |       |       |
|-------|-----------|-------|-------|
|       | Female (Hz) | Male (Hz) |
|       | F0 | F1 | F2 | F3 | F0 | F1 | F2 | F3 |
| a     |     |     |     |     |     |     |     |     |
| Singleton | 195 | 1050 | 1540 | 2902 | 116 | 769 | 1298 | 2489 |
| Geminate | 28 | 79 | 120 | 182 | 3 | 10 | 58 | 98 |
| i     |     |     |     |     |     |     |     |     |
| Singleton | 194 | 1025 | 1546 | 2894 | 122 | 774 | 1304 | 2538 |
| Geminate | 30 | 57 | 122 | 197 | 9 | 25 | 65 | 136 |
| u     |     |     |     |     |     |     |     |     |
| Singleton | 216 | 414 | 2770 | 3511 | 184 | 295 | 2237 | 2956 |
| Geminate | 26 | 54 | 107 | 181 | 41 | 15 | 60 | 133 |

Table XVI - Average and standard deviation of F0, F1, F2 and F3 in reference frames V1 CENTER, V1 OFFSET and V1-TO-C TRANSITION for liquids, for female vs. male speakers, averaged over repetitions, speakers and consonants (values are in Hz).
Table XVII - Average and standard deviation of pitch F0 and formants F1, F2 and F3 in reference frames V2 ONSET and V2 CENTER, and of pitch F0 in reference frames C1 ONSET, C CENTER and C OFFSET for words containing liquids, for female vs. male speakers, averaged over repetitions, speakers and consonants (frequencies are in Hz).

A multi-factor univariate ANOVA test was then performed using Form, Vowel and Consonant as fixed factors. Results are presented in Table XVIII, where a checked cell indicates a significant difference between average values for the parameter with respect to factor. Table XVIII shows that gemination only led to statistically significant variations for frequency domain parameters for female speakers, in particular for F1 and F3 at V1 OFFSET, for F2 at C ONSET, and again for F1 at C OFFSET. In the case of male speakers, gemination never led to significant variations of any parameter in any frame.

Vowel was the only factor leading to significant differences in F1, F2 and F3 for both female and male speakers in most frames, with the exception of F3 at C OFFSET for male speakers and at V1 OFFSET for female speakers.

### C ONSET / C CENTER / C OFFSET

|        | Female (Hz) |       | Male (Hz) |       |
|--------|-------------|-------|-----------|-------|
|        | F0  | F1  | F2  | F3  | F0  | F1  | F2  | F3  |
| a      | Singleton | Mean | 187 | 184 | 187 | 108 | 106 | 108 |
|        |        | Std  | 26  | 23  | 26  | 4   | 9   | 10  |
|        | Geminate| Mean | 195 | 190 | 189 | 116 | 106 | 100 |
|        |        | Std  | 29  | 24  | 24  | 9   | 16  | 11  |
| i      | Singleton| Mean | 206 | 221 | 231 | 148 | 122 | 110 |
|        |        | Std  | 26  | 50  | 48  | 55  | 26  | 14  |
|        | Geminate| Mean | 208 | 209 | 191 | 130 | 115 | 122 |
|        |        | Std  | 33  | 23  | 25  | 33  | 18  | 47  |
| u      | Singleton| Mean | 213 | 205 | 210 | 142 | 112 | 108 |
|        |        | Std  | 85  | 81  | 82  | 44  | 8   | 9   |
|        | Geminate| Mean | 218 | 208 | 208 | 151 | 122 | 120 |
|        |        | Std  | 26  | 23  | 24  | 45  | 26  | 26  |

### V2 ONSET

|        | Female (Hz) |       | Male (Hz) |       |
|--------|-------------|-------|-----------|-------|
|        | F0  | F1  | F2  | F3  | F0  | F1  | F2  | F3  |
| a      | Singleton | Mean | 186 | 764 | 1688 | 2912 | 105 | 604 | 1350 | 2497 |
|        |        | Std  | 22  | 85  | 140  | 312  | 10  | 50  | 105  | 90   |
|        | Geminate| Mean | 187 | 707 | 1582 | 2779 | 103 | 577 | 1287 | 2497 |
|        |        | Std  | 24  | 85  | 94   | 397  | 13  | 79  | 55   | 139  |
| i      | Singleton| Mean | 209 | 381 | 2524 | 3140 | 152 | 314 | 1966 | 2549 |
|        |        | Std  | 25  | 41  | 97   | 83   | 75  | 30  | 82   | 98   |
|        | Geminate| Mean | 197 | 410 | 2432 | 3077 | 126 | 329 | 1901 | 2534 |
|        |        | Std  | 25  | 27  | 185  | 191  | 31  | 42  | 76   | 119  |
| u      | Singleton| Mean | 198 | 407 | 985  | 2542 | 107 | 353 | 967  | 2039 |
|        |        | Std  | 79  | 157 | 65   | 227  | 9   | 29  | 54   | 165  |
|        | Geminate| Mean | 210 | 425 | 1037 | 2226 | 130 | 347 | 998  | 2011 |
|        |        | Std  | 18  | 39  | 92   | 211  | 49  | 30  | 126  | 157  |

### V2 CENTER

|        | Female (Hz) |       | Male (Hz) |       |
|--------|-------------|-------|-----------|-------|
|        | F0  | F1  | F2  | F3  | F0  | F1  | F2  | F3  |
| a      | Singleton | Mean | 187 | 978 | 1562 | 2947 | 106 | 734 | 1357 | 2473 |
|        |        | Std  | 10  | 63  | 97   | 295  | 11  | 26  | 94   | 100  |
|        | Geminate| Mean | 185 | 942 | 1535 | 2945 | 104 | 735 | 1370 | 2490 |
|        |        | Std  | 19  | 67  | 71   | 346  | 13  | 18  | 98   | 201  |
| i      | Singleton| Mean | 211 | 373 | 2751 | 3331 | 164 | 302 | 2205 | 2790 |
|        |        | Std  | 12  | 30  | 96   | 162  | 64  | 26  | 100  | 96   |
|        | Geminate| Mean | 202 | 388 | 2718 | 3282 | 130 | 321 | 2125 | 2694 |
|        |        | Std  | 15  | 24  | 131  | 188  | 25  | 38  | 119  | 94   |
| u      | Singleton| Mean | 199 | 407 | 852  | 2883 | 129 | 382 | 823  | 2263 |
|        |        | Std  | 79  | 157 | 20   | 200  | 43  | 27  | 50   | 187  |
|        | Geminate| Mean | 202 | 412 | 852  | 2849 | 135 | 346 | 864  | 2285 |
|        |        | Std  | 28  | 34  | 40   | 179  | 36  | 33  | 77   | 222  |
Vowel also led to significant variations of F0 for male speakers in all frames except for C ONSET, C CENTER and V2 ONSET. Consonant led to significant differences in F1 in consonant-related frames, in particular at C ONSET (males only), C CENTER and C OFFSET (both female and male speakers), and sporadically in other parameters for male speakers: F0 at C ONSET, C OFFSET and V2 CENTER, F2 at C CENTER, and F3 at V1 OFFSET, V1-TO-C TRANSITION and V2 ONSET.

As a general comment, data for female speakers showed a lower impact of all factors on each parameter. In particular, F0 was not significantly influenced by any factor in any frame.

|                | Female  |          |          | Male    |          |          |
|----------------|---------|----------|----------|---------|----------|----------|
|                |         | F0       | F1       | F2      | F3       | F0       | F1       | F2     |
| V1 CENTER      | Form    | X        | X        | X       | X        | X        | X        |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | N/A     | N/A      | N/A     | N/A      | N/A      | N/A      |        |
| V1 OFFSET      | Form    | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | X        | X        | X        | X        | X        | X        |        |
| V1-TO-C TRANSITION | Form | N/A | N/A | N/A | N/A | N/A | N/A |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | X        | X        | X        | X        | X        | X        |        |
| C ONSET        | Form    | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
|                | Vowel   | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
|                | Consonant | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
| C CENTER       | Form    | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | X        | X        | X        | X        | X        | X        |        |
| C OFFSET       | Form    | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | X        | X        | X        | X        | X        | X        |        |
| V2 ONSET       | Form    | X        | X        | X       | X        | X        | X        |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | X        | X        | X        | X        | X        | X        |        |
| V2 CENTER      | Form    | N/A      | N/A      | N/A     | N/A      | N/A      | N/A      |        |
|                | Vowel   | X        | X        | X       | X        | X        | X        |        |
|                | Consonant | X        | X        | X        | X        | X        | X        |        |

Table XVIII – Results of the multi-factor univariate ANOVA test performed on frequency domain parameters in reference frames defined in Section 3.1.1 for words containing liquids using Form, Vowel and Consonant as fixed factors; a checked cell indicates a significant difference of average values for the parameter with respect to the factor.

4.2.3. Results in the energy domain

Table XIX shows mean values and standard deviations for energy domain parameters for each combination of vowels [a, i, u], consonants [l, r] and forms (singleton vs. geminate), averaged over speakers and repetitions. A direct inspection of data in Table XIX does not highlight any clear trend for any of the parameters, in particular in relation to the gemination.

Following the same approach adopted for nasals, a multi-factor univariate ANOVA test considering the fixed factors Form, Vowel, Consonant and Gender was performed over all combined words. Results are presented in Table XX, and show that Form is typically not a significant factor, since only the EtotC parameter shows significant variation with gemination. As for the other factors, Vowel is, by far, the one leading to a stronger impact, since it leads to significant variations of all energy-related parameters. Gender and Consonant only led to sporadic significant differences, respectively for EtotC and EiV1cent (Gender) and EiV1-C and EiCcen (Consonant).
Finally, speakers, and only in the frequency domain $F_0$ with $\theta_m$, but not $\theta_{ml}$. Male speakers showed a significant difference between average values for the parameter with respect to the factor.

Form, Vowel, Consonant and Gender as fixed factors for all words; a checked cell at the intersection between a parameter and a factor indicates a significant difference between average values for the parameter with respect to the factor.

5. Discussion

5.1. Effect of gemination in nasals

Results of the analysis presented in Section 4.1.1 showed a significant increase in consonant duration and a decrease of pre-consonant vowel duration for all combinations of vowels and consonants, and for both female and male speakers. No significant variation was observed in the post-consonant vowel duration. Word duration Utd was only marginally affected by gemination, with significant variations observed for all combinations of vowels with [m], but not with [n], for which only coarticulation with [i] led to significant Utd variations.

In the frequency domain $F_0$ significantly increased when moving from singleton to geminate only for male speakers, and only for reference frames related to V1, in particular in words containing vowels [i] and [u]. No significant variations were observed for formants in any frame for neither female nor male speakers.

Finally, the total energy of the consonant $E_{ac}$ showed significant variations with gemination, while all the other energy domain parameters were not affected by gemination.
5.2. Effect of gemination in liquids

Time domain parameters for liquids were strongly correlated with gemination. Cd, V1d and Utd were in fact significantly different in singletons vs. geminates for all combinations of vowels and consonants, although the impact on Utd was typically weaker, as shown by higher p values when compared to V1d and even more to Cd. The analysis of frequency domain parameters was carried out for liquids by studying both pitch F0 and formants F1, F2 and F3 in vowel frames and F0 in consonant frames.

Finally, in analogy with results observed for nasals, the total energy of the consonant $E_{tot}$ was the only parameter showing significant variation with gemination.

5.3. Comparison in terms of acoustic correlates of gemination in nasals and liquids

Results of the present study on nasals and liquids are in agreement with the results presented for stops in (Esposito and Di Benedetto, 1999): a significantly increased consonant duration, and a corresponding reduced pre-consonant vowel length in geminates. Results on frequency and energy parameters show instead a negligible and sporadic effect of gemination on such parameters. The comparison of the impact of gemination in nasals vs. liquids is therefore carried out on temporal parameters only. This analysis will be extended in (Di Benedetto and De Nardis, 2019) to affricates, fricatives and will also include the previously analysed stop category.

Table XXI summarizes mean values and standard deviations for liquids and nasals, averaged over all repetitions, speakers, consonants and vowels. Table XXI shows that consonant duration Cd is the parameter with largest relative variation across all consonant categories (+133% in nasals, +187% in liquids) followed by pre-consonant vowel duration V1d (+32% in nasals, +41% in liquids).

Table XXI - Mean values and standard deviations of the time related parameters averaged over all the repetitions, speakers, consonants and vowels for nasals and liquids.

| Nasals | V1d | Cd   | V2d   | Utd   | Cd/V1d |
|--------|-----|------|-------|-------|--------|
|        | Mean| 185.52| 90.64 | 130.05| 404.20 | 0.51   |
|        | StD | 27.45 | 14.14 | 25.43 | 45.07  | 0.12   |
|        | Mean| 124.56| 211.75| 124.25| 460.57 | 1.77   |
|        | StD | 20.95 | 33.33 | 25.43 | 43.02  | 0.56   |
|        | Mean| 171.92| 60.56 | 100.21| 384.1  | 0.36   |
|        | StD | 25.75 | 15.33 | 22.1  | 40.53  | 0.11   |
|        | Mean| 121.81| 174.2 | 87.74 | 443.86 | 1.52   |
|        | StD | 27.54 | 28.69 | 21.45 | 42.87  | 0.51   |

Results of the analysis on the significance of time domain parameter variations for nasals (Table IV) and liquids (Table XIII) are in good agreement with the analysis carried out in (Esposito and Di Benedetto 1999) for stops. Comparison in terms of Spearman Rank correlation shows that both nasals and liquids present a high negative correlation between V1d and Cd (< -0.65), while a weaker correlation is observed when the analysis is restricted to geminate words, and no correlation at all is present when only singleton words are considered.

5.4. Classification of geminate vs. singleton words in nasals and liquids

Results presented in Section 4 highlighted that only time domain parameters are consistently and significantly affected by gemination. Time domain parameters were thus adopted as test variables for Maximum Likelihood Classification tests (Dillon and Goldstein, 1984) of geminate vs. singleton words. Table XXII shows the classification percentage error for tests on nasals and liquids using V1d, Cd and V2d for male and female speakers and for all words combined. Results in Table XXII are in good agreement with the results of the ANOVA tests shown in Section 4; Cd, that is the parameter that presented the most significant variations with gemination also led to the lowest classification error rates. Classification tests using V1d led to higher error percentages, coherently with the weaker significance for V1d variations observed in Section 4. Additional tests were carried out, to investigate the combination of multiple parameters in the classification of geminate vs. singleton words. The analysis focused on the combination of Cd and V1d. Parameters were combined in two ways. First, they were used as variables in a bidimensional MLC test, following the same approach adopted in (Esposito and Di Benedetto, 1999) for stops. Secondly, the ratio Cd/V1d was used in a unidimensional test, following what suggested in (Pickett et al. 1999).
Table XXII - Percentage of singleton vs. geminate classification errors for nasal and liquid consonants based on unidimensional MLC tests on time domain parameters V1d, Cd, and V2d for separate female and male speakers, and for all combined words.

Table XXIII shows the classification error percentage for the following three cases: 1) female speakers, 2) male speakers and 3) all speakers combined. Results of the bidimensional tests indicate that in nasals the introduction of V1d allowed to remove the residual classification errors observed in Table XXII when only Cd was used. In liquids the classification based on Cd was already error free, and the introduction of V1d did not affect the classification performance. The results of the unidimensional test using the Cd/V1d ratio does not consistently lead to improved classification rate. In nasals a slight improvement was observed for male speakers when switching from C1d to C1d/V1d, while classification rate did not change for combined speakers, and actually degraded from perfect classification to a 1.9% error rate for female speakers. In liquids a small classification rate loss was observed in all groups: 0.5% for combined speakers, 0.9% for both male and female speakers.

Table XXIII - Percentage of singleton vs. geminate classification errors for nasal and liquid consonants in a bidimensional test using (Cd, V1d), and in an unidimensional MLC test using the Cd/V1d ratio for separate female and male speakers, and for all combined words.

The thresholds on Cd/V1d that led to the best classification performance in the MLC test, corresponding to the Points of Equal Probability (PEPs) between the two Gaussian distributions fitted on singleton vs. geminate data, are presented in Table XXIV. Table XXIV also presents the thresholds that led to the best classification performance in a heuristic test that explored all possible thresholds, without assuming Gaussian distributions for singletons vs. geminates; the heuristic test was motivated by the limited size of the set of words, that might not be properly fitted by a Gaussian distribution. Table XXIV shows that the best classification performance was obtained with a threshold in the order of 0.75 in most cases for both consonant classes, the only exception being words including liquids pronounced by female speakers, for which the threshold was below 0.6.

Table XXIV – Thresholds for singleton vs. geminate classification in nasal and liquid consonants using the Cd/V1d ratio for separate female and male speakers, and for all combined words; thresholds were determined both as the Point of Equal Probability (PEP) resulting from the assumption of Gaussian distributions for the two groups of geminate and singleton words, and heuristically as the value that minimizes the number of classification errors.
These results can be compared with those presented in (Pickett et al. 1999) for classification of singleton vs. geminate stop consonants. Pickett et al. (1999) found in fact by visual inspection of Cd and V1d values that classification based on Cd/V1d with an arbitrary value of 1 led to satisfactory classification error rates across different speaking rates, indicating an invariance property of Cd/V1d with speaking rate. One might thus wonder whether Cd/V1d shows a similar invariance property across different consonant categories. Results in Table XXIV for nasals and liquids seem indeed to indicate that Cd/V1d may show some form of invariance across consonants, at least in terms of best classification threshold, although our threshold is lower than the one proposed in (Pickett et al. 1999) for stops. In order to better assess this aspect, Figure 4 presents the classification error rate in the heuristic test as a function of the Cd/V1d threshold for combined male and female speakers data for nasals and liquids.

![Figure 4](image)

**Figure 4** - Classification error rate as a function of the Cd/V1d threshold for combined male and female speakers data for nasals vs. liquids.

Results in Figure 4 show that for both categories the error rate remains close to its minimum value for a wide range of Cd/V1d threshold values, due to the clear separation between singleton and separate words in terms of Cd/V1d; the value of 1 proposed in (Pickett et al. 1999) falls within this range for nasals but not for liquids, leaving the question of a common threshold existing for all consonant categories open. The potential role of Cd/V1d as an across-consonant classification parameter is further investigated in (Di Benedetto, M. G., and De Nardis, L. 2019) where a similar analysis is carried out for affricates and fricatives, and the use of this parameter for classifying geminate vs. singleton consonants of all consonant classes, stops, nasals, liquids, fricatives, and affricates is tested.

6. Conclusions

This work investigated the impact of gemination on nasal and liquid Italian consonants, based on acoustic analyses of disyllabic words (VCV vs. VCCV) in a symmetrical context of cardinal vowels [a, i, u]. These words belong to the GEMMA project database (GEMMA, 2019). Time domain, frequency domain and energy domain measurements were collected in different frames within the word, corresponding to crucial events such as vowel-to-consonant transition and vowel and consonant stable portions.

The most relevant outcomes can be summarized as follows:

- a general tendency of shortening the pre-consonant vowel and of lengthening the consonant in a geminate word, that was observed for stops in previous studies (Esposito and Di Benedetto, 1999), was confirmed for both nasals and liquids;
- a careful examination of the speech materials under study highlighted a high degree of correlation between the two aforementioned effects when considering the full set of singletons vs. geminates. A
weaker correlation is already present in geminates vs. geminates, while no correlation was observed in singletons vs. singletons. This result is important since it quantifies a hypothesis suggested by Shrotriya et al., (1995), that the observed effect is related to a need of preserving rhythmical structures;

- the analysis of energy-related parameters highlighted that the energy of the consonant Exc was significantly affected by gemination for both nasals and liquids. This result marks a clear difference with stops, for which no significant variations in energy parameters were observed (Esposito and Di Benedetto, 2019);
- the use of the primary acoustic cue Cd for classification of singletons vs. geminates led to the best classification rates for both nasals and liquids. In the case of nasals, error-free classification was obtained using Cd, while in liquids residual classification errors were eliminated by combining the primary cue with first vowel duration V1d in a bidimensional classifier;
- the Cd/V1d ratio was investigated as an across-consonant parameter for detecting gemination; satisfactory classification rates were obtained in both nasals and liquids and stops using a same threshold value. This threshold value was however different from the one proposed in previous studies for classification of gemination in stops (Pickett et al. 1999), questioning the invariance of Cd/V1d with consonant category. A further discussion on this aspect will be included in the companion paper by considering all five consonant categories.

7. References
Abramson A. S. (1998). “The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length,” in Warotamasiikkadit, U. and Panakul, T. (eds.), Papers from the Fourth Annual Meeting of the Southeast Asian Linguistics Society, 1994. Tempe, Arizona: Program for Southeast Asian Studies, Arizona State University.
Abramson A. S. (1999). “Fundamental frequency as a cue to word-initial consonant length: Pattani Malay,” Proceedings of ICPhS99, San Francisco, pp. 591-594.
Al-Tamimi, J., and Khattab, G. (2011). “Multiple cues for the singleton-geminate contrast in Lebanese Arabic: Acoustic investigation of stops and fricatives,” in Proceedings of the 17th ICPhS, Hong Kong, China, pp. 212-215.
Al-Tamimi, J., and Khattab, G. (2015). “Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants,” The Journal of the Acoustical Society of America, Vol. 138, Issue 1, pp. 344-360, 2015. DOI: 10.1121/1.4922514.
Argiolas F., Macri F., and Di Benedetto M. G. (1995). “Acoustic analysis of Italian [r] and [l],” Journal of the Acoustical Society of America vol. 97, no. 5, pt.2, p. 3418, May 1995. DOI: 10.1121/1.412455.
Arvaniti A. (1999). “Effects of speaking rate on the timing of single and geminate sonorants,” in Proceedings of ICPhS99, San Francisco, pp 599-602.
Arvaniti, A., and Tserdanelis, G. (2000). “On the phonetics of geminates: Evidence from Cypriot Greek,” in Proceedings of ICSLP, Beijing, China, pp. 559–562.
Canepari L. (1992). Manuale di pronuncia italiana, Zanichelli.
Cohn A.C., Ham W.H. and Podesva R. J. (1999). “The phonetic realization of singleton-geminate contrasts in three languages of Indonesia,” Proceedings of ICPhS99, San Francisco, pp 587-590.
Di Benedetto, M. G. (2000). “Gemination in Italian: the GEMMA project,” Journal of the Acoustical Society of America Vol. 108, Issue 5, pt. 2, p. 2507, November 2000. DOI: 10.1121/1.4743263.
Di Benedetto, M. G., and De Nardis, L. (2019). “Consonant gemination in Italian: the affricate and fricative case,” submitted to Speech Communication.
Dillon W.R. and Goldstein M. (1984). Multivariate Analysis. Wiley J. & Sons.
Esposito, A., and Di Benedetto, M. G. (1999). “Acoustic and Perceptual Study of Gemination in Italian Stops,” The Journal of the Acoustical Society of America, ASA, Vol. 30, pp. 175-185, 1999. DOI: 10.1121/1.428056.
GEMMA (2019). GEMMA project webpage, available at: http://newyork.ing.uniroma1.it/project_gemma.php.
Giovanardi M. and Di Benedetto M.G. (1998). “Acoustic analysis of singleton and geminate fricatives in Italian,” European Student journal of Language and Speech WEB-SLS, Vol. 1998, pp. 1-13.
Hansen, B. B. (2004). “Production of Persian geminate stops: Effects of varying speaking rate,” in Proceedings of the 2003 Texas Linguistics Society Conference, edited by A. Agwuele, W. Warren, and S.-H. Park (Cascadilla Proceedings Project, Somerville, MA), pp. 86–95.
Hassan, Z. M. (2003). “Temporal compensation between vowel and consonant in Swedish and Arabic in sequences of CV:C and CVC: and the word overall duration,” in Proceedings of PHONUM, Sweden, Vol. 9, pp. 45–48.

Hirata, Y., and Whiton, J. (2005). “Effects of speaking rate on the single/geminate stop distinction in Japanese,” The Journal of the Acoustical Society of America, Vol. 118, Issue 13, pp. 1647-1660. DOI: 10.1121/1.2000807.

Idemaru, K., and Guion, S. G. (2008). “Acoustic covariants of length contrast in Japanese stops,” Journal of the International Phonetic Association, Vol. 38, Issue 2, pp. 167–186.

Khattab, G., and Al-Tamimi, J. (2014). “Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure,” Lab. Phon. Vol. 5, Issue 2, pp. 231–269.

Local J. and Simpson A.P. (1999). “Phonological contrast and phonetic realization: the case of Berber stops,” Proceedings of 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 WEB-SLS, Vol. 2000, pp. 1-11.

Muljacic D. (1972). Fonologia della lingua italiana. Il Mulino.

Pickett E.R., Blumstein S.E. and Burton M.W. (1999). “Effects of speaking rate on Singleton/Geminate consonant contrast in Italian,” Phonetica, vol. 56, pp. 135–157. DOI:10.1159/000028448.

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, Stockholm, 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 Electronics Engineering, Sapienza University of Rome, Italy.

Rossetti R. (1994). “Gemination of Italian stops,” Journal of the Acoustical Society of America 95, 2 SP25, p. 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 ICPhS95, Stockholm, pp.132-135.

Tserdanelis, G., and Arvaniti, A. (2001). “The acoustic characteristics of geminate consonants in Cypriot Greek,” in Proceedings of the Fourth International Conference on Greek Linguistics, Thessaloniki, Greece (University Studio Press S.A., Thessaloniki, Greece), pp. 29–36.

Vecsys (2019). Vecsys, now a division of Bertin IT http://www.bertin-it.com.

Acknowledgments

This work was partly funded by Sapienza University of Rome within research projects (“ex-quota 60%”, “ricerca di Facoltà”) in the years 1991-2019 and supported in part by the Radcliffe Institute for Advanced Study at Harvard University. The authors wish to thank Marco Mattei, Federico Macri and Francesca Argiolas for their collaboration on the GEMMA project while they were interns at the Speech Lab at the DIET Department working toward their Master of Science degree in Electrical Engineering.